

# Second-order numerical schemes for decoupled forward-backward stochastic differential equations with jumps

Weidong Zhao

*Department of Mathematics, Shandong University, Jinan, Shandong 250100, China*

*Email: wdzhao@sdu.edu.cn*

Wei Zhang

*Department of Mathematics, Beijing University of Technology, Beijing 100022, China*

*Email: weizhang0313@bjut.edu.cn*

Guannan Zhang

*Department of Computational and Applied Mathematics, Oak Ridge National Laboratory, Oak Ridge,*

*Tennessee 37831, United States*

*Email: zhangg@ornl.gov*

## Abstract

We propose new numerical schemes for decoupled forward-backward stochastic differential equations (FBSDEs) with jumps, where the stochastic dynamics are driven by a  $d$ -dimensional Brownian motion and an independent compensated Poisson random measure. A semi-discrete scheme is developed for discrete time approximation, which is constituted by a classic scheme for the forward SDE [17, 25] and a novel scheme for the backward SDE. Under some reasonable regularity conditions, we prove that the semi-discrete scheme can achieve second-order convergence in approximating the FBSDEs of interest; and such convergence rate does not require jump-adapted temporal discretization. Next, to add in spatial discretization, a fully discrete scheme is developed by designing accurate quadrature rules for estimating the involved conditional mathematical expectations. Several numerical examples are given to illustrate the effectiveness and the high accuracy of the proposed schemes.

*Mathematics subject classification:* 60H35, 60H10, 65C20, 65C30

*Key words:* Decoupled FBSDEs with Lévy jumps, backward Kolmogorov equation, non-linear Feynman-Kac formula, second-order convergence, error estimates.

## 1. Introduction

In this work, we study numerical solution of decoupled forward-backward stochastic differential equations (FBSDEs) with jumps, where the underlying stochastic jump processes are characterized by Poisson random measures. The term “decoupled” refers to the fact that the forward SDE is independent of the solution of the backward SDE. This work is motivated by a wide variety of applications offered by FBSDEs. In finance and insurance, FBSDEs-based approaches [23, 29] have gained a great attention by both academics and practitioners, because FBSDEs provide us a unified framework to describe the mathematical problems which arise in option pricing [13], portfolio hedging [14], market utility maximization [2] and risk measures [24, 26], etc. Moreover, in the presence of jump behaviors in many financial problems [25], Lévy jump processes have been incorporated into FBSDEs [11, 14], so as to accurately capture

and properly interpret event-driven stochastic phenomena, such as corporate defaults, operational failures, insured events, etc. In mathematics, one can relate FBSDEs with jumps to a class of nonlinear partial integro-differential equations (PIDEs), based on the extension of the nonlinear Feynman-Kac theory studied in [1]. As such, FBSDEs become a powerful probabilistic technique for studying analytical and numerical solutions and properties of the PIDEs, where the nonlocal integral operators of the PIDEs are characterized by Poisson random measures in the framework of FBSDEs. In engineering science, a particular application of the PIDEs is to model anomalous diffusion [19], i.e., super-diffusion and sub-diffusion, that has been verified experimentally to be present in various applications, e.g., contaminant transport in groundwater and plasma physics. In this setting, FBSDEs-based probabilistic numerical schemes have been developed in [31] to solve the governing PIDEs, which illustrated effectiveness of the FBSDEs models.

There are many theoretical results on FBSDEs with jumps over the past two decades. The existence and uniqueness were proved by Tang and Li [29] for backward stochastic differential equations with Poisson jumps and Lipschitzian coefficients, which was then extended, by Rong in [27], to the case of non-Lipschitzian coefficients. In [1], Barles, Buckdahn and Pardoux established a comparison theorem for decoupled FBSDEs with jumps as well as the link between such FBSDEs and PIDEs, which generalized the results in [21, 22] to the case of a natural filtration associated with a Brownian motion and a Poisson random measure. After that, in the context of FBSDEs with jumps, Øksendal and Sulem [20] established maximum principles, and Royer [28] introduced nonlinear expectations. For a general overview of related topics, see [9, 11] and the references therein.

The obstacle of applying FBSDEs with jumps to real-world engineering and finance problems results from the challenge of solving FBSDEs analytically or numerically. Since it is typically difficult to obtain analytical solutions, numerical solutions are highly desired in practical applications. Numerical methods for FBSDEs without jumps have been well studied in the literature [3, 7, 8, 10, 12, 15, 32–34, 36], nevertheless, there are very few numerical schemes developed for FBSDEs with jumps, and most of those schemes only focused on temporal discretization. For instance, a Picard’s iterative method was provided in [18], and numerical schemes of backward SDE were studied in [4, 5]. Due to the aforementioned applications of FBSDEs with jumps, it is of great significance to develop high-order temporal-spatial discretization schemes for solving not only the FBSDEs but also the PIDEs and related engineering problems.

In this paper, we propose novel numerical schemes for decoupled FBSDEs driven by a  $d$ -dimensional Brownian motion and an independent compensated Poisson random measure. In general, the approximation of the FBSDEs under consideration includes two steps, i.e., constructing a semi-discrete scheme for temporal discretization, and extending it to a fully discrete scheme by incorporating effective spatial discretization. By imposing appropriate regularity conditions on the coefficients, the generator and the terminal condition, we rigorously prove the *second-order* convergence of the semi-discrete scheme with respect to  $\Delta t$ . In spatial discretization, a carefully designed quadrature rule is critical to approximate all the involved conditional mathematical expectations which are, in this case, multiple integrals with respect to both the Brownian motion and the Poisson random measure. The integrals with respect to the Brownian motion is estimated by the Gauss-Hermite rule. For the integrals with respect to the Poisson random measure, we propose a general quadrature rule for the case that the jump component has *finite* activities. A specific form of the quadrature rule can be determined based on the type of the underlying Lévy measure. For the numerical experiments in §6, the Lévy measure

is defined as a uniform distribution on bounded domains, so that Gauss-Legendre rule is an appropriate choice. Moreover, to avoid the explosion of the total number of quadrature points with the increase of time steps, we construct a piecewise Lagrange interpolating polynomials on a pre-determined spatial mesh, which are used to evaluate the integrand at all quadrature points.

The main contributions of this paper are as follows:

- propose a second-order discrete time approximation (semi-discrete) scheme for decoupled FBSDEs with jumps.
- rigorously analyze the convergence rate of the proposed discrete time approximation scheme with respect to  $\Delta t$ .
- propose a fully discrete scheme by developing new quadrature rules for estimating involved conditional mathematical expectations.

The outline of the paper is organized as follows. In §2, we introduce the mathematical description of the FBSDEs under consideration. In §3, we propose the semi-discrete scheme, i.e., time discretization, for the FBSDEs of interest. Rigorous error analysis for the proposed semi-discrete scheme is conducted in §4. The fully discrete scheme for the case of Poisson random measures with finite activities is proposed in §5. Numerical examples are given in §6, to show the effectiveness and the high accuracy of our approach. Finally, several concluding remarks and discussions about our future work are given in §7.

## 2. Preliminaries

Let  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbb{P})$  be a stochastic basis satisfying the usual hypotheses of completeness, i.e.,  $\mathcal{F}_0$  contains all the sets of  $\mathbb{P}$ -measure zero and possesses right continuity, i.e.,  $\mathcal{F}_t = \mathcal{F}_{t+}$ . The filtration  $\{\mathcal{F}_t\}_{0 \leq t \leq T}$  is assumed to be generated by two mutually independent processes, i.e., one  $d$ -dimensional Brownian motion  $W_t = (W_t^1, \dots, W_t^d)^\top$  and one Poisson random measure  $\mu(A, t)$  on  $E \times [0, T]$  where  $E = \mathbb{R}^q \setminus \{0\}$  is equipped with its Borel field  $\mathcal{E}$ . The compensator of  $\mu$  and the resulting compensated Poisson random measure are denoted by  $\nu(de, dt) = \lambda(de)dt$  and  $\tilde{\mu}(de, dt) = \mu(de, dt) - \lambda(de)dt$ , respectively, such that  $\{\tilde{\mu}(A \times [0, t]) = (\mu - \nu)(A \times [0, t])\}_{0 \leq t \leq T}$  is a martingale for all  $A \in \mathcal{E}$ .  $\lambda(de)$  is assumed to be a  $\sigma$ -finite measure on  $(E, \mathcal{E})$  satisfying

$$\int_E (1 \wedge |e|^2) \lambda(de) < +\infty,$$

where  $|\cdot|$  denotes the standard Euclidean norm in Euclidean spaces.

In the probability space  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbb{P})$ , we introduce the following forward-backward stochastic differential equation with jumps

$$\begin{cases} X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s + \int_0^t \int_E c(s, X_{s-}, e) \tilde{\mu}(de, ds), \\ Y_t = \xi + \int_t^T f(s, X_s, Y_s, Z_s, \Gamma_s) ds - \int_t^T Z_s dW_s - \int_t^T \int_E U_s(e) \tilde{\mu}(de, ds), \end{cases} \quad (2.1)$$

where the quadruplet  $(X_t, Y_t, Z_t, U_t)$  is the unknown,  $b : [0, T] \times \mathbb{R}^q \rightarrow \mathbb{R}^q$  is referred to as the drift coefficient,  $\sigma : [0, T] \times \mathbb{R}^q \rightarrow \mathbb{R}^{q \times d}$  is referred to as the local diffusion coefficient,

$c : [0, T] \times \mathbb{R}^q \times E \rightarrow \mathbb{R}^q$  is referred to as the jump coefficient,  $f : [0, T] \times \mathbb{R}^q \times \mathbb{R}^p \times \mathbb{R}^{p \times d} \times \mathbb{R}^p \rightarrow \mathbb{R}^p$  is referred to as the generator of the FBSDE, and the process  $\Gamma_s$  is defined by  $\Gamma_s = \int_E U_s(e) \eta(e) \lambda(de)$  for a given bounded function  $\eta : E \rightarrow \mathbb{R}$ , i.e.,  $\sup_{e \in E} |\eta(e)| < +\infty$ . The terminal condition  $\xi$  is an  $\mathcal{F}_T$ -measurable random vector in  $\mathbb{R}^q$ . A quadruplet  $(X_t, Y_t, Z_t, U_t)$  is called an  $L^2$ -adapted solution if it is an  $\{\mathcal{F}_t\}$ -adapted, square integrable processes satisfying the FBSDEs in (2.1).

Under standard assumptions on the given data  $b, \sigma, f, \varphi$  and  $c$  (see [1] for details), there exists a unique solution  $(Y_t, Z_t, U_t) \in S^2 \times L^2(W) \times L^2(\tilde{\mu})$  for the backward SDE in (2.1), where  $S^2$  is the set of  $\{\mathcal{F}_t\}$ -adapted càdlàg processes  $\{Y_t, 0 \leq t \leq T\}$  such that  $\|Y\|_{S^2}^2 := \mathbb{E} \left[ \left( \sup_{0 \leq t \leq T} |Y_t| \right)^2 \right] < \infty$ ,  $L^2(W)$  the set of  $\mathcal{F}_t$ -progressively measurable  $q \times d$  dimensional processes  $\{Z_t, 0 \leq t \leq T\}$  such that  $\|Z\|_{L^2(W)}^2 := \mathbb{E} \left[ \int_0^T |Z_t|^2 dt \right] < \infty$ , and  $L^2(\tilde{\mu})$  the set of mappings  $U : \Omega \times [0, T] \times E \rightarrow \mathbb{R}$  such that  $\|U\|_{L^2(\tilde{\mu})}^2 := \mathbb{E} \left[ \int_0^T \int_E U_t(e)^2 \lambda(de) dt \right] < \infty$ .

Now we introduce a class of *nonlinear* partial integro-differential equations (PIDEs) that will be related to the FBSDEs in (2.1) later. We consider the unique viscosity solution  $u(t, x) \in \mathcal{C}([0, T] \times \mathbb{R}^q)$  of the following nonlinear PIDE, i.e.,

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + \tilde{\mathcal{L}}[u](t, x) + f(t, x, u, \sigma \nabla u, \mathcal{B}[u]) = 0, & \text{for } (t, x) \in [0, T] \times \mathbb{R}^q, \\ u(T, x) = \varphi(x), & \text{for } x \in \mathbb{R}^q, \end{cases} \quad (2.2)$$

where  $\varphi(x)$  is the terminal condition at the time  $t = T$ ,  $\tilde{\mathcal{L}}$  is the second-order integral-differential operator of the form

$$\begin{aligned} \tilde{\mathcal{L}}[u](t, x) = & \sum_{i=1}^q b_i(t, x) \frac{\partial u}{\partial x^i}(t, x) + \frac{1}{2} \sum_{i,j=1}^q (\sigma \sigma^\top)_{i,j}(t, x) \frac{\partial^2 u}{\partial x^i \partial x^j}(t, x) \\ & + \int_E \left( u(t, x + c(t, x, e)) - u(t, x) - \sum_{i=1}^q \frac{\partial u}{\partial x^i}(t, x) c(t, x, e) \right) \lambda(de), \end{aligned} \quad (2.3)$$

and  $\mathcal{B}$  is an integral operator defined as

$$\mathcal{B}[u](t, x) = \int_E [u(t, x + c(t, x, e)) - u(t, x)] \eta(e) \lambda(de).$$

For  $(t, x) \in [0, T] \times \mathbb{R}^q$ , let  $\mathbb{E}_t^x[\cdot]$  denote the mathematical expectation under the condition that  $X_t = x$ , i.e.,  $\mathbb{E}_t^x[\cdot] := \mathbb{E}[\cdot | X_t = x]$ . To relate the FBSDEs in (2.1) with the PIDE in (2.2), we consider the FBSDEs of the following form

$$\begin{cases} X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x}) dr + \int_t^s \sigma(r, X_r^{t,x}) dW_r + \int_t^s \int_E c(r, X_{r-}^{t,x}, e) \tilde{\mu}(de, dr), \\ Y_s^{t,x} = \xi + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}, \Gamma_r^{t,x}) dr - \int_s^T Z_r^{t,x} dW_r - \int_s^T \int_E U_r^{t,x}(e) \tilde{\mu}(de, dr), \end{cases} \quad (2.4)$$

where the solution is  $(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x}, U_s^{t,x})$  and  $\Gamma_s^{t,x} = \int_E U_s^{t,x}(e) \eta(e) \lambda(de)$  for  $t \leq s \leq T$ . Note that the superscripts in (2.4) indicate the fact that the forward SDE in (2.4) starts from the time-space point  $(t, x) \in [0, T] \times \mathbb{R}^q$ .

According to Theorem 3.4 in [1], if the terminal condition  $\xi$  of the FBSDEs is a function

of  $X_T^{t,x}$ , defined by  $\xi = \varphi(X_T^{t,x})$  ( $\varphi(\cdot)$  is the terminal condition of the PIDE), then the triple  $(Y_s^{t,x}, Z_s^{t,x}, U_s^{t,x})$  for  $t \leq s \leq T$  can be represented by the unique viscosity solution  $u(t, x)$  of the PIDE (2.2) as follows:

$$\begin{cases} Y_s^{t,x} = u(s, X_s^{t,x}), \\ Z_s^{t,x} = \sigma(s, X_s^{t,x}) \nabla u(s, X_s^{t,x}), \\ U_s^{t,x} = u(s, X_{s-}^{t,x} + c(s-, X_{s-}^{t,x}, e)) - u(s, X_{s-}^{t,x}), \end{cases} \quad (2.5)$$

where  $\nabla u$  denotes the gradient of  $u$  with respect to  $x$  and the function  $\Gamma_s^{t,x}$  is defined by  $\Gamma_s^{t,x} = \mathcal{B}[u](s, X_s^{t,x})$ . Particularly, when  $s = t$ , we have  $u(t, x) = Y_t^{t,x} = \mathbb{E}[Y_t | X_t = x]$ .

### 3. The semi-discrete scheme for FBSDEs with jumps

In this section, we propose a numerical scheme for discrete-time approximation of the FBSDEs under consideration. Instead of the FBSDEs (2.1), we will use the conditional representation of the FBSDEs given in (2.4) throughout this section. Specifically, discretizations of the forward SDE and backward SDE are discussed in §3.1 and §3.2, respectively, and the main numerical scheme is proposed in §3.3. To proceed, we introduce the following time partition for the interval  $[0, T]$ :

$$\mathcal{T} := \{0 = t_0 < t_1 < \dots < t_N = T\} \quad (3.1)$$

with  $\Delta t_n := t_{n+1} - t_n$  and  $\Delta t := \max_{0 \leq n \leq N-1} \Delta t_n$ . We assume that the time partition  $\mathcal{T}$  has the following regularity:

$$\frac{\max_{0 \leq n \leq N-1} \Delta t_n}{\min_{0 \leq n \leq N-1} \Delta t_n} \leq c_0, \quad (3.2)$$

where  $c_0 \geq 1$  is a real positive constant. We remark that  $\mathcal{T}$  is *not* a jump-adapted partition.

#### 3.1. Discretization of the forward SDE

Due to the decoupling of the FBSDEs in (2.4), the forward SDE can be discretized separately. Here we briefly recall some classic numerical schemes and their properties discussed in [25]. Any of these schemes can serve as the approximation of the forward SDE in our schemes for the FBSDE. By setting  $t = t_n$ ,  $s = t_{n+1}$  and  $x = X^n$  in (2.4), the forward SDE can be written as

$$\begin{aligned} X_{t_{n+1}}^{t_n, X^n} = & X^n + \int_{t_n}^{t_{n+1}} b(s, X_s^{t_n, X^n}) ds + \int_{t_n}^{t_{n+1}} \sigma(s, X_s^{t_n, X^n}) dW_s \\ & + \int_{t_n}^{t_{n+1}} \int_E c(s, X_{s-}^{t_n, X^n}, e) \tilde{\mu}(de, ds), \end{aligned} \quad (3.3)$$

where we assume that the solution  $X_s^{t_n, X^n}$  starts at the time instant  $t = t_n$  and spatial location  $X_{t_n}^{t_n, X^n} = X^n$ . By using the Itô-Taylor expansion, numerical schemes of strong-order  $\beta$  (or the weak-order  $\beta$ ) [25] can be represented in a general form, i.e.,

$$X^{n+1} = X^n + \Phi(t_n, t_{n+1}, X^n, I_{\mathcal{J} \in \mathcal{A}_\beta}), \quad (3.4)$$

where  $\Phi$  is the incremental,  $\mathcal{A}_\beta$  is a hierarchical set such that the convergence rate of the scheme is  $\beta$  in a strong or weak sense. Details of the index set  $I_{\mathcal{J} \in \mathcal{A}_\beta}$  and the definition of  $\mathcal{A}_\beta$  can be found in [25] (pp. 196 and pp. 290). The scheme (3.4) has the following properties:

- *Stability*: for an integer  $r > 0$ , there exists a constant  $C \in (0, \infty)$  such that

$$\max_{0 \leq n \leq N} \mathbb{E}[|X^n|^r] \leq C(1 + \mathbb{E}[|X_0|^r]). \quad (3.5)$$

- *Approximation error*: there exist positive real numbers  $r_1, r_2, r_3, \alpha, \beta, \gamma$  such that for any function  $g \in C_P^{2\beta+2}$ , we have

$$\begin{aligned} \left| \mathbb{E}_{t_n}^{X^n} [g(X_{t_{n+1}}^{t_n, X^n}) - g(X^{n+1})] \right| &\leq C(1 + |X^n|^{2r_1})(\Delta t)^{\beta+1}, \\ \left| \mathbb{E}_{t_n}^{X^n} [(g(X_{t_{n+1}}^{t_n, X^n}) - g(X^{n+1})) \Delta \tilde{W}_{t_{n+1}}^\top] \right| &\leq C(1 + |X^n|^{2r_2})(\Delta t)^{\gamma+1}, \\ \left| \mathbb{E}_{t_n}^{X^n} [(g(X_{t_{n+1}}^{t_n, X^n}) - g(X^{n+1})) \Delta \tilde{\mu}_{t_{n+1}}^*] \right| &\leq C(1 + |X^n|^{2r_3})(\Delta t)^{\alpha+1}, \end{aligned} \quad (3.6)$$

where  $C_P^{2\beta+2}$  is the set of  $2\beta+2$  times continuously differentiable functions which, together with their derivatives of order up to  $2\beta+2$ , have at most polynomial growth. According to Theorem 6.4.1 and Theorem 12.3.4 in [25], it is easy to derive that  $\alpha = \beta = \gamma$  for strong and weak Taylor schemes. In this paper, we prove in Theorem 4.2 that the second-order convergence of the proposed semi-discrete scheme for the FBSDE in (2.1) requires  $\alpha = \beta = \gamma = 2$ .

### 3.2. Discretization of the backward SDE

Now we study the discretization of the backward SDE in (2.4) driven by the process  $X_s^{t_n, X^n}$  in (3.3) for  $s \in [t_n, t_{n+1}]$ . Within the interval  $[t_n, t_{n+1}]$ , the backward SDE can be rewritten as

$$Y_{t_n}^{t_n, X^n} = Y_{t_{n+1}}^{t_n, X^n} + \int_{t_n}^{t_{n+1}} f_s^{t_n, X^n} ds - \int_{t_n}^{t_{n+1}} Z_s^{t_n, X^n} dW_s - \int_{t_n}^{t_{n+1}} \int_E U_s^{t_n, X^n}(e) \tilde{\mu}(de, ds), \quad (3.7)$$

where  $f_s^{t_n, X^n}$  denotes  $f(s, X_s^{t_n, X^n}, Y_s^{t_n, X^n}, Z_s^{t_n, X^n}, \Gamma_s^{t_n, X^n})$  for notational simplicity. Due to the relation between  $\Gamma_s^{t, x}$  and  $U_s^{t, x}$ , in what follows, all the numerical schemes for the backward SDE will be proposed to approximate  $(Y_s^{t_n, X^n}, Z_s^{t_n, X^n}, \Gamma_s^{t_n, X^n})$ . Since there are three unknown stochastic processes involved in (3.7), we now construct three discretized reference equations for  $Y_{t_n}^{t_n, X^n}$ ,  $Z_{t_n}^{t_n, X^n}$  and  $\Gamma_{t_n}^{t_n, X^n}$  in §3.2.1, §3.2.2 and §3.2.3, respectively, which are the foundation of the formal semi-discrete scheme discussed in §3.3 for the FBSDEs.

#### 3.2.1. The reference equation for $Y_{t_n}^{t_n, X^n}$

Taking the conditional mathematical expectation  $\mathbb{E}_{t_n}^{X^n}[\cdot]$  on both sides of (3.7), we obtain

$$Y_{t_n}^{t_n, X^n} = \mathbb{E}_{t_n}^{X^n} [Y_{t_{n+1}}^{t_n, X^n}] + \int_{t_n}^{t_{n+1}} \mathbb{E}_{t_n}^{X^n} [f_s^{t_n, X^n}] ds, \quad (3.8)$$

due to the fact that  $\int_{t_n}^{t_{n+1}} Z_s^{t_n, X^n} dW_s$  and  $\int_{t_n}^{t_{n+1}} \int_E U_s^{t_n, X^n}(e) \tilde{\mu}(de, ds)$  for  $t > t_n$  are martingales. Note that the integrand  $\mathbb{E}_{t_n}^{X^n} [f_s^{t_n, X^n}]$  is a *deterministic* function of  $s \in [t_n, t_{n+1}]$  under the  $\sigma$ -algebra  $\mathcal{F}_{t_n}$ . Thus, numerical integration approaches can be used to approximate the temporal

integral in (3.8). In this effort, we use the Crank-Nicolson scheme, i.e., the trapezoidal rule, such that

$$\int_{t_n}^{t_{n+1}} \mathbb{E}_{t_n}^{X^n} [f_s^{t_n, X^n}] ds = \frac{1}{2} \Delta t_n f_{t_n}^{t_n, X^n} + \frac{1}{2} \Delta t_n \mathbb{E}_{t_n}^{X^n} [f_{t_{n+1}}^{t_n, X^n}] + R_y^n, \quad (3.9)$$

where the residual  $R_y^n$  is

$$R_y^n := \int_{t_n}^{t_{n+1}} \left\{ \mathbb{E}_{t_n}^{X^n} [f_s^{t_n, X^n}] - \frac{1}{2} f_{t_n}^{t_n, X^n} - \frac{1}{2} \mathbb{E}_{t_n}^{X^n} [f_{t_{n+1}}^{t_n, X^n}] \right\} ds. \quad (3.10)$$

Substituting (3.9) into (3.8), we obtain the reference equation for solving  $Y_{t_n}^{t_n, X^n}$ :

$$Y_{t_n}^{t_n, X^n} = \mathbb{E}_{t_n}^{X^n} [Y_{t_{n+1}}^{t_n, X^n}] + \frac{1}{2} \Delta t_n f_{t_n}^{t_n, X^n} + \frac{1}{2} \Delta t_n \mathbb{E}_{t_n}^{X^n} [f_{t_{n+1}}^{t_n, X^n}] + R_y^n. \quad (3.11)$$

### 3.2.2. The reference equation for $Z_{t_n}^{t_n, X^n}$

To proceed, we introduce a new Gaussian process  $\Delta \tilde{W}_s$  defined by

$$\Delta \tilde{W}_s = 2\Delta W_s - \frac{3}{\Delta t_n} \int_{t_n}^s (r - t_n) dW_r, \quad \forall s \in [t_n, t_{n+1}], \quad (3.12)$$

where  $\Delta W_s = W_s - W_{t_n}$  is the  $d$ -dimensional standard Brownian motion in the FBSDEs in (2.4). It is easy to see that  $\Delta \tilde{W}_s = (\Delta \tilde{W}_s^1, \Delta \tilde{W}_s^2, \dots, \Delta \tilde{W}_s^d)^\top$  is also a  $d$ -dimensional Gaussian process with the properties  $\mathbb{E}_{t_n}^{X^n} [\Delta \tilde{W}_s] = 0$ ,  $\mathbb{E}_{t_n}^{X^n} [\Delta \tilde{W}_s^i \Delta \tilde{W}_s^j] = 0$  for  $i \neq j$ , and

$$\begin{aligned} \mathbb{E}_{t_n}^{X^n} [(\Delta \tilde{W}_s^i)^2] &= \mathbb{E}_{t_n}^{X^n} \left[ \left( 2\Delta W_s^i - \frac{3}{\Delta t_n} \int_{t_n}^s (r - t_n) dW_r^i \right)^2 \right] \\ &= 4(s - t_n) - \frac{6(s - t_n)^2}{\Delta t_n} + \frac{3(s - t_n)^3}{\Delta t_n^2}, \quad \text{for } i = 1, \dots, d. \end{aligned}$$

In the case of  $s = t_{n+1}$ , we have  $\mathbb{E}_{t_n}^{X^n} [\Delta \tilde{W}_{t_{n+1}}^i] = 0$  and  $\mathbb{E}_{t_n}^{X^n} [(\Delta \tilde{W}_{t_{n+1}}^i)^2] = \Delta t_n$  for  $i = 1, \dots, d$ .

Multiplying (3.7) by the transpose of  $\Delta \tilde{W}_{t_{n+1}}$  in (3.12), and taking the conditional mathematical expectation  $\mathbb{E}_{t_n}^{X^n} [\cdot]$  on both sides, we obtain

$$\begin{aligned} 0 &= \mathbb{E}_{t_n}^{X^n} \left[ Y_{t_{n+1}}^{t_n, X^n} \Delta \tilde{W}_{t_{n+1}}^\top \right] + \int_{t_n}^{t_{n+1}} \mathbb{E}_{t_n}^{X^n} \left[ f_s^{t_n, X^n} \Delta \tilde{W}_{t_{n+1}}^\top \right] ds \\ &\quad - \mathbb{E}_{t_n}^{X^n} \left[ \int_{t_n}^{t_{n+1}} Z_s^{t_n, X^n} dW_s \cdot \Delta \tilde{W}_{t_{n+1}}^\top \right]. \end{aligned} \quad (3.13)$$

Then, the right endpoint rule is used to discretize the first temporal integral in (3.13), such that

$$\int_{t_n}^{t_{n+1}} \mathbb{E}_{t_n}^{X^n} \left[ f_s^{t_n, X^n} \Delta \tilde{W}_{t_{n+1}}^\top \right] ds = \Delta t_n \mathbb{E}_{t_n}^{X^n} \left[ f_{t_{n+1}}^{t_n, X^n} \Delta \tilde{W}_{t_{n+1}}^\top \right] + R_{z,1}^n, \quad (3.14)$$

where  $R_{z,1}^n := \int_{t_n}^{t_{n+1}} \mathbb{E}_{t_n}^{X^n} [f_s^{t_n, X^n} \Delta \tilde{W}_{t_{n+1}}^\top] ds - \Delta t_n \mathbb{E}_{t_n}^{X^n} [f_{t_{n+1}}^{t_n, X^n} \Delta \tilde{W}_{t_{n+1}}^\top]$  is the residual. For the

second temporal integral in (3.13), based on the properties of  $\Delta \tilde{W}_{t_{n+1}}^\top$ , we discretize it by

$$-\mathbb{E}_{t_n}^{X^n} \left[ \int_{t_n}^{t_{n+1}} Z_s^{t_n, X^n} dW_s \cdot \Delta \tilde{W}_{t_{n+1}}^\top \right] = -\frac{1}{2} \Delta t_n Z_{t_n}^{t_n, X^n} + R_{z,2}^n, \quad (3.15)$$

where the residual is  $R_{z,2}^n := \frac{1}{2} \Delta t_n Z_{t_n}^{t_n, X^n} - \mathbb{E}_{t_n}^{X^n} [\int_{t_n}^{t_{n+1}} Z_s^{t_n, X^n} dW_s \cdot \Delta \tilde{W}_{t_{n+1}}^\top]$ . Substituting (3.14) and (3.15) into (3.13), we obtain the reference equation for  $Z_{t_n}^{t_n, X^n}$ , i.e.,

$$\frac{1}{2} \Delta t_n Z_{t_n}^{t_n, X^n} = \mathbb{E}_{t_n}^{X^n} \left[ Y_{t_{n+1}}^{t_n, X^n} \Delta \tilde{W}_{t_{n+1}}^\top \right] + \Delta t_n \mathbb{E}_{t_n}^{X^n} \left[ f_{t_{n+1}}^{t_n, X^n} \Delta \tilde{W}_{t_{n+1}}^\top \right] + R_z^n, \quad (3.16)$$

where  $R_z^n := R_{z,1}^n + R_{z,2}^n$ .

### 3.2.3. The reference equation for $\Gamma_{t_n}^{t_n, X^n}$

Similar to the definition of  $\Delta \tilde{W}_s$ , by using the compensated Poisson random measure  $\tilde{\mu}(de, ds)$  in (2.4), we define a new stochastic process  $\Delta \tilde{\mu}_s^*$  as

$$\Delta \tilde{\mu}_s^* = \int_{t_n}^s \int_E \left( 2 - \frac{3(t - t_n)}{\Delta t_n} \right) \eta(e) \tilde{\mu}(de, dt), \quad \forall s \in [t_n, t_{n+1}]. \quad (3.17)$$

Then, multiplying (3.7) by  $\Delta \tilde{\mu}_{t_{n+1}}^*$  and taking the conditional mathematical expectation  $\mathbb{E}_{t_n}^{X^n}[\cdot]$  on both sides, we obtain

$$\begin{aligned} 0 &= \mathbb{E}_{t_n}^{X^n} \left[ Y_{t_{n+1}}^{t_n, X^n} \Delta \tilde{\mu}_{t_{n+1}}^* \right] + \int_{t_n}^{t_{n+1}} \mathbb{E}_{t_n}^{X^n} \left[ f_s^{t_n, X^n} \Delta \tilde{\mu}_{t_{n+1}}^* \right] ds \\ &\quad - \mathbb{E}_{t_n}^{X^n} \left[ \int_{t_n}^{t_{n+1}} \int_E U_s^{t_n, X^n}(e) \tilde{\mu}(de, ds) \Delta \tilde{\mu}_{t_{n+1}}^* \right]. \end{aligned} \quad (3.18)$$

Analogous to the reference equation (3.16), we also discretize the first temporal integral in (3.18) using the right endpoint rule, such that

$$\int_{t_n}^{t_{n+1}} \mathbb{E}_{t_n}^{X^n} \left[ f_s^{t_n, X^n} \Delta \tilde{\mu}_{t_{n+1}}^* \right] ds = \Delta t_n \mathbb{E}_{t_n}^{X^n} \left[ f_{t_{n+1}}^{t_n, X^n} \Delta \tilde{\mu}_{t_{n+1}}^* \right] + R_{\Gamma,1}^n, \quad (3.19)$$

where  $R_{\Gamma,1}^n := \int_{t_n}^{t_{n+1}} \mathbb{E}_{t_n}^{X^n} [f_s^{t_n, X^n} \Delta \tilde{\mu}_{t_{n+1}}^*] ds - \Delta t_n \mathbb{E}_{t_n}^{X^n} [f_{t_{n+1}}^{t_n, X^n} \Delta \tilde{\mu}_{t_{n+1}}^*]$  is the residual. For the second temporal integral in (3.18), we have

$$\begin{aligned} & - \mathbb{E}_{t_n}^{X^n} \left[ \int_{t_n}^{t_{n+1}} \int_E U_s^{t_n, X^n}(e) \tilde{\mu}(de, ds) \Delta \tilde{\mu}_{t_{n+1}}^* \right] \\ &= - \mathbb{E}_{t_n}^{X^n} \left[ \int_{t_n}^{t_{n+1}} \int_E U_{t_n}^{t_n, X^n}(e) \tilde{\mu}(de, ds) \Delta \tilde{\mu}_{t_{n+1}}^* \right] + R_{\Gamma,2}^n \\ &= - \mathbb{E}_{t_n}^{X^n} \left[ \Gamma_{t_n}^{t_n, X^n} \int_{t_n}^{t_{n+1}} \left( 2 - \frac{3(s - t_n)}{\Delta t_n} \right) ds \right] + R_{\Gamma,2}^n \\ &= - \frac{1}{2} \Delta t_n \Gamma_{t_n}^{t_n, X^n} + R_{\Gamma,2}^n, \end{aligned} \quad (3.20)$$



where  $R_{\Gamma,2}^n := \frac{1}{2}\Delta t_n \Gamma_{t_n}^{t_n, X^n} - \mathbb{E}_{t_n}^{X^n} [\int_{t_n}^{t_{n+1}} \int_E U_s^{t_n, X^n}(e) \tilde{\mu}(de, ds) \Delta \tilde{\mu}_{t_{n+1}}^*]$ . By (3.18), (3.19) and (3.20), we obtain the reference equation for  $\Gamma_{t_n}^{t_n, X^n}$ , i.e.,

$$\frac{1}{2}\Delta t_n \Gamma_{t_n}^{t_n, X^n} = \mathbb{E}_{t_n}^{X^n} [Y_{t_{n+1}}^{t_n, X^n} \Delta \tilde{\mu}_{t_{n+1}}^*] + \Delta t_n \mathbb{E}_{t_n}^{X^n} [f_{t_{n+1}}^{t_n, X^n} \Delta \tilde{\mu}_{t_{n+1}}^*] + R_{\Gamma}^n, \quad (3.21)$$

where  $R_{\Gamma}^n := R_{\Gamma,1}^n + R_{\Gamma,2}^n$ .

### 3.3. The semi-discrete scheme

Now we combine the approximation in (3.4) of the forward SDE and the three reference equations in (3.11), (3.16) and (3.21) to propose our semi-discrete scheme (temporal discretization) for the FBSDEs in (2.1). Let  $(X^{n+1}, Y^n, Z^n, \Gamma^n)$  denote the approximation to the exact solution  $(X_{t_{n+1}}^{t_n, X^n}, Y_{t_n}^{t_n, X^n}, Z_{t_n}^{t_n, X^n}, \Gamma_{t_n}^{t_n, X^n})$  of the FBSDEs in (2.1) for  $n = N-1, \dots, 0$ . Based on the partition  $\mathcal{T}$  of the time interval  $[0, T]$ , the approximate solution  $(X^{n+1}, Y^n, Z^n, \Gamma^n)$  is constructed following the procedure in Scheme 1.

**Scheme 1.** Given the initial condition  $X_0$  for the forward SDE in (2.1) and the terminal condition  $(Y^N, Z^N, \Gamma^N)$  for the backward SDE in (2.1), solve the approximate solution  $(X^{n+1}, Y^n, Z^n, \Gamma^n)$ , for  $n = N-1, \dots, 0$ , by

$$X^{n+1} = X^n + \Phi(t_n, t_{n+1}, X^n, I_{\mathcal{J} \in \mathcal{A}_\beta}), \quad (3.22)$$

$$\frac{1}{2}\Delta t_n Z^n = \mathbb{E}_{t_n}^{X^n} [Y^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top] + \Delta t_n \mathbb{E}_{t_n}^{X^n} [f^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top], \quad (3.23)$$

$$\frac{1}{2}\Delta t_n \Gamma^n = \mathbb{E}_{t_n}^{X^n} [Y^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*] + \Delta t_n \mathbb{E}_{t_n}^{X^n} [f^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*], \quad (3.24)$$

$$Y^n = \mathbb{E}_{t_n}^{X^n} [Y^{n+1}] + \frac{1}{2}\Delta t_n f^n + \frac{1}{2}\Delta t_n \mathbb{E}_{t_n}^{X^n} [f^{n+1}], \quad (3.25)$$

where  $f^{n+1} = f(t_{n+1}, X^{n+1}, Y^{n+1}, Z^{n+1}, \Gamma^{n+1})$ ,  $f^n = (t_n, X^n, Y^n, Z^n, \Gamma^n)$ ,  $\Delta \tilde{W}_{t_{n+1}}$  and  $\Delta \tilde{\mu}_{t_{n+1}}^*$  are defined according to (3.12) and (3.17) by setting  $s = t_{n+1}$ , respectively.

From the dependence of (3.22)–(3.25), we can see that the scheme in (3.22) for  $X^{n+1}$  is independent of the other three schemes, so that, at each time step,  $X^{n+1}$  is always firstly determined. Then, by observing that (3.23) and (3.24) are *explicit* schemes, we can solve  $Z^n$  and  $\Gamma^n$  by substituting  $X^{n+1}$  into (3.23) and (3.24), respectively. Next, since (3.25) includes  $f^n$  that depends on  $Y^n$ ,  $Z^n$  and  $\Gamma^n$ , it is an *implicit* scheme for  $Y^n$ . If the generator  $f^n$  is *nonlinear* and Lipschitz continuous with respect to  $Y^n$ , then  $Y^n$  can be obtained by substituting  $X^{n+1}$ ,  $Z^n$  and  $\Gamma^n$  into (3.25) and solving a nonlinear equation.

In addition, we would like to discuss the application of Scheme 1 to discrete time approximations of the PIDE in (2.2) when the terminal condition  $\xi$  of the FBSDEs is a function of  $X_T$ , i.e.,  $\xi = \varphi(X_T)$ . The goal is to construct an approximate solution  $u(t_n, x)$  for  $n = N-1, \dots, 1, 0$  and  $x \in \mathbb{R}^q$ . Specifically, based on the relation between  $u(t, x)$  and  $(X_t, Y_t, Z_t, \Gamma_t)$  in (2.5), the discrete time approximation, denoted by  $u^n(x)$ , is defined by

$$u^n(x) := \mathbb{E}[Y^n | X^n = x] \approx u(t_n, x) = \mathbb{E}[Y_{t_n} | X_{t_n} = x] \quad \text{for } x \in \mathbb{R}^q. \quad (3.26)$$

It is easy to see that both  $Y_{t_n}$  and  $Y^n$  are *deterministic* values under the conditions  $X_{t_n} = x$  and  $X^n = x$ , respectively. Moreover, the convergence of  $(X^{n+1}, Y^n, Z^n, \Gamma^n)$  to  $(X_{t_{n+1}}, Y_{t_n}, Z_{t_n}, \Gamma_{t_n})$

as  $N \rightarrow \infty$  will ensure the convergence of  $u^n(x)$  to  $u(t_n, x)$ . Hence, Scheme 1 can be viewed as an effective probabilistic scheme for the PIDE in (2.2). Moreover,  $Z^n$  and  $\Gamma^n$  provide approximations of  $\sigma \nabla u$  and  $\mathcal{B}[u]$  which enable accurate characterization of local and nonlocal diffusive fluxes in practical engineering problems.

#### 4. Error estimates for the semi-discrete scheme

In this section, we estimate the truncation error of Scheme 1. Since error estimates for the scheme for the forward SDEs have been well established in the literature (see [25] and the references therein), we focus on analyzing the approximation error of  $(Y^n, Z^n, \Gamma^n)$  for  $n = 0, \dots, N-1$ . The general procedure of our analysis is similar to that for classic time-stepping schemes. We first construct an upper bound of the global truncation error of  $(Y^n, Z^n, \Gamma^n)$  by recursively accumulating local truncation errors. Then, we estimate all the local truncation errors in the upper bound, which relates the global truncation error to the maximum time step size  $\Delta t := \max_{1 \leq n \leq N} \Delta t_n$ .

To proceed, we need to specify the definition of the approximation error of  $(Y^i, Z^i, \Gamma^i)$ . For  $i = 0, \dots, N$ , the errors of  $Y^i$ ,  $Z^i$ ,  $\Gamma^i$  and  $f^i := f(t_i, X^i, Y^i, Z^i, \Gamma^i)$  are respectively defined by

$$\begin{aligned} e_y^i &:= Y_{t_i}^{t_i, X^i} - Y^i, & e_z^i &:= Z_{t_i}^{t_i, X^i} - Z^i, \\ e_\Gamma^i &:= \Gamma_{t_i}^{t_i, X^i} - \Gamma^i, & e_f^i &:= f_{t_i}^{t_i, X^i} - f^i, \end{aligned} \quad (4.1)$$

where  $Y_{t_i}^{t_i, X^i} := \mathbb{E}[Y_{t_i} | X_{t_i} = X^i]$  and likewise for  $Z_{t_i}^{t_i, X^i}$ ,  $\Gamma_{t_i}^{t_i, X^i}$ ,  $f_{t_i}^{t_i, X^i}$ . It should be noted that  $Y_{t_{i+1}}^{t_i, X^i}$  and  $Y_{t_{i+1}}^{t_{i+1}, X^{i+1}}$  for  $0 \leq i \leq N$  are usually different stochastic processes because of the difference between  $X_{t_{i+1}}^{t_i, X^i}$  and  $X^{i+1}$ . This fact can be easily shown with the use of the solution  $u(t, x)$  of the PIDE in (2.2). According to the relationship in (2.5), it is easy to see that

$$Y_{t_{i+1}}^{t_i, X^i} = u(t_{i+1}, X_{t_{i+1}}^{t_i, X^i}), \quad Y_{t_{i+1}}^{t_{i+1}, X^{i+1}} = u(t_{i+1}, X^{i+1}),$$

where  $X_{t_{i+1}}^{t_i, X^i}$  and  $X^{i+1}$  are obtained by (3.3) and (3.22), respectively. As such, we introduce the following residual notations that will be used later:

$$\begin{aligned} \mathcal{R}_{y_1}^i &:= \mathbb{E}_{t_i}^{X^i} \left[ Y_{t_{i+1}}^{t_i, X^i} - Y_{t_{i+1}}^{t_{i+1}, X^{i+1}} \right], \\ \mathcal{R}_{y_2}^i &:= \mathbb{E}_{t_i}^{X^i} \left[ f_{t_{i+1}}^{t_i, X^i} - f_{t_{i+1}}^{t_{i+1}, X^{i+1}} \right], \\ \mathcal{R}_{z_1}^i &:= \mathbb{E}_{t_i}^{X^i} \left[ \left( Y_{t_{i+1}}^{t_i, X^i} - Y_{t_{i+1}}^{t_{i+1}, X^{i+1}} \right) \Delta \tilde{W}_{t_{i+1}}^\top \right], \\ \mathcal{R}_{z_2}^i &:= \mathbb{E}_{t_i}^{X^i} \left[ \left( f_{t_{i+1}}^{t_i, X^i} - f_{t_{i+1}}^{t_{i+1}, X^{i+1}} \right) \Delta \tilde{W}_{t_{i+1}}^\top \right], \\ \mathcal{R}_{\Gamma_1}^i &:= \mathbb{E}_{t_i}^{X^i} \left[ \left( Y_{t_{i+1}}^{t_i, X^i} - Y_{t_{i+1}}^{t_{i+1}, X^{i+1}} \right) \Delta \tilde{\mu}_{t_{i+1}}^* \right], \\ \mathcal{R}_{\Gamma_2}^i &:= \mathbb{E}_{t_i}^{X^i} \left[ \left( f_{t_{i+1}}^{t_i, X^i} - f_{t_{i+1}}^{t_{i+1}, X^{i+1}} \right) \Delta \tilde{\mu}_{t_{i+1}}^* \right] \end{aligned} \quad (4.2)$$

for  $i = 0, \dots, N-1$ . Note that the above residuals represent the local weak approximations of the scheme (3.22) for solving forward SDEs.

In the following theorem, we construct an upper bound of the errors  $e_y^n$ ,  $e_z^n$  and  $e_\Gamma^n$  ( $0 \leq n \leq N-1$ ) with the use of the residuals  $\mathcal{R}_{y_1}^i$ ,  $\mathcal{R}_{z_1}^i$  and  $\mathcal{R}_{\Gamma_1}^i$  in (3.11), (3.16) and (3.21), respectively,

for  $i = n, \dots, N$ , as well as the residuals defined in (4.2).

**Theorem 4.1.** *Based on the partition  $\mathcal{T}$  in (3.1) of the time interval  $[0, T]$ , if the generator  $f(t, x, y, z, \gamma)$  is Lipschitz continuous with respect to  $x, y, z$  and  $\gamma$  where Lipschitz constant is denoted by  $L$ , then with sufficiently small time step  $\Delta t := \max_{1 \leq i \leq N} \Delta t_i$ , the errors  $e_y^n$ ,  $e_z^n$  and  $e_\Gamma^n$  in (4.1) for  $n = 0, \dots, N - 1$  can be bounded by*

$$\begin{aligned}
& \mathbb{E}[|e_y^n|^2] + \Delta t \sum_{i=n}^{N-1} \left( \frac{1+C\Delta t}{1-C\Delta t} \right)^{i-n} \mathbb{E}[|e_z^i|^2 + |e_\Gamma^i|^2] \\
& \leq C' \left( \mathbb{E}[|e_y^N|^2] + \Delta t \mathbb{E}[|e_z^N|^2 + |e_\Gamma^N|^2] \right) \\
& \quad + \sum_{i=n}^{N-1} \left( \frac{1+C\Delta t}{1-C\Delta t} \right)^{i-n} \frac{C}{1-C\Delta t} \mathbb{E} \left[ \frac{1}{\Delta t} (|R_y^i|^2 + |R_z^i|^2 + |R_\Gamma^i|^2) \right. \\
& \quad \left. + \frac{1}{\Delta t} (|\mathcal{R}_{y_1}^i|^2 + |\mathcal{R}_{z_1}^i|^2 + |\mathcal{R}_{\Gamma_1}^i|^2) + \Delta t (|\mathcal{R}_{y_2}^i|^2 + |\mathcal{R}_{z_2}^i|^2 + |\mathcal{R}_{\Gamma_2}^i|^2) \right], \tag{4.3}
\end{aligned}$$

where  $C$  is a positive constant depending on  $L$  and  $c_0$  defined in (3.2),  $C'$  is a positive constant depending on  $c_0, T$  and  $L$ , the residuals  $R_y^i, R_z^i$  and  $R_\Gamma^i$  for  $i = n, \dots, N$  are defined in (3.11), (3.16) and (3.21), respectively, and  $\mathcal{R}_{y_1}^i, \mathcal{R}_{y_2}^i, \mathcal{R}_{z_1}^i, \mathcal{R}_{z_2}^i, \mathcal{R}_{\Gamma_1}^i, \mathcal{R}_{\Gamma_2}^i$  are defined in (4.2).

*Proof.* This proof consists of four steps. Step 1, 2 and 3 are dedicated to the estimating  $e_y^n$ ,  $e_z^n$  and  $e_\Gamma^n$ , respectively, and those estimates are combined together at Step 4 that completes the proof.

• **Step 1:** Estimating the error  $e_y^n = Y_{t_n}^{t_n, X^n} - Y^n$ .

Subtracting the scheme (3.25) from the reference equation (3.11), we have

$$\begin{aligned}
e_y^n &= \mathbb{E}_{t_n}^{X^n} \left[ Y_{t_{n+1}}^{t_n, X^n} - Y^{n+1} \right] + \frac{1}{2} \Delta t_n \left( f_{t_n}^{t_n, X^n} - f^n \right) \\
& \quad + \frac{1}{2} \Delta t_n \mathbb{E}_{t_n}^{X^n} \left[ f_{t_{n+1}}^{t_n, X^n} - f^{n+1} \right] + R_y^n \\
&= \mathbb{E}_{t_n}^{X^n} \left[ Y_{t_{n+1}}^{t_n, X^n} - Y_{t_{n+1}}^{t_{n+1}, X^{n+1}} + Y_{t_{n+1}}^{t_{n+1}, X^{n+1}} - Y^{n+1} \right] + \frac{\Delta t_n}{2} \left( f_{t_n}^{t_n, X^n} - f^n \right) \\
& \quad + \frac{\Delta t_n}{2} \mathbb{E}_{t_n}^{X^n} \left[ f_{t_{n+1}}^{t_n, X^n} - f_{t_{n+1}}^{t_{n+1}, X^{n+1}} + f_{t_{n+1}}^{t_{n+1}, X^{n+1}} - f^{n+1} \right] + R_y^n \\
&= \mathbb{E}_{t_n}^{X^n} [e_y^{n+1}] + \frac{\Delta t_n}{2} e_f^n + \frac{\Delta t_n}{2} \mathbb{E}_{t_n}^{X^n} [e_f^{n+1}] + \mathcal{R}_{y_1}^n + \frac{\Delta t_n}{2} \mathcal{R}_{y_2}^n + R_y^n.
\end{aligned}$$

Then under the conditions of the theorem, we have the estimate

$$\begin{aligned}
|e_y^n| &\leq \left| \mathbb{E}_{t_n}^{X^n} [e_y^{n+1}] \right| + \frac{\Delta t_n}{2} \left| \mathbb{E}_{t_n}^{X^n} [e_f^{n+1}] \right| + \frac{\Delta t_n}{2} |e_f^n| + |\mathcal{R}_{y_1}^n| + \frac{\Delta t_n}{2} |\mathcal{R}_{y_2}^n| + |R_y^n| \\
&\leq \left| \mathbb{E}_{t_n}^{X^n} [e_y^{n+1}] \right| + \frac{\Delta t_n}{2} L \mathbb{E}_{t_n}^{X^n} [|e_y^{n+1}| + |e_z^{n+1}| + |e_\Gamma^{n+1}|] \\
&\quad + \frac{\Delta t_n}{2} L (|e_y^n| + |e_z^n| + |e_\Gamma^n|) + |\mathcal{R}_{y_1}^n| + \frac{\Delta t_n}{2} |\mathcal{R}_{y_2}^n| + |R_y^n|.
\end{aligned}$$

For the squared error  $|e_y^n|^2$ , given any positive real number  $\gamma$  and positive integer  $m$ , by using the inequalities  $(a+b)^2 \leq (1+\gamma\Delta t)a^2 + (1+\frac{1}{\gamma\Delta t})b^2$  and  $(\sum_{n=1}^m a_n)^2 \leq m \sum_{n=1}^m a_n^2$ , we have

the following estimate.

$$\begin{aligned}
|e_y^n|^2 &\leq (1 + \gamma\Delta t) \left| \mathbb{E}_{t_n}^{X^n} [e_y^{n+1}] \right|^2 + \left( 1 + \frac{1}{\gamma\Delta t} \right) \left\{ \frac{\Delta t_n}{2} L (|e_y^n| + |e_z^n| + |e_\Gamma^n|) \right. \\
&\quad \left. + \frac{\Delta t_n}{2} L \mathbb{E}_{t_n}^{X^n} [|e_y^{n+1}| + |e_z^{n+1}| + |e_\Gamma^{n+1}|] + |\mathcal{R}_{y_1}^n| + \frac{\Delta t_n}{2} |\mathcal{R}_{y_2}^n| + |R_y^n| \right\}^2 \\
&\leq (1 + \gamma\Delta t) \left| \mathbb{E}_{t_n}^{X^n} [e_y^{n+1}] \right|^2 + \left\{ \frac{15\Delta t^2 L^2}{4} (|e_y^n|^2 + |e_z^n|^2 + |e_\Gamma^n|^2) \right. \\
&\quad + \frac{15\Delta t^2 L^2}{4} \mathbb{E}_{t_n}^{X^n} [|e_y^{n+1}|^2 + |e_z^{n+1}|^2 + |e_\Gamma^{n+1}|^2] + 5|\mathcal{R}_{y_1}^n|^2 \\
&\quad + \frac{5\Delta t^2}{4} |\mathcal{R}_{y_2}^n|^2 + 5|R_y^n|^2 \left. \right\} + \frac{1}{\gamma} \left\{ \frac{15\Delta t L^2}{4} (|e_y^n|^2 + |e_z^n|^2 + |e_\Gamma^n|^2) \right. \\
&\quad + \frac{15\Delta t L^2}{4} \mathbb{E}_{t_n}^{X^n} [|e_y^{n+1}|^2 + |e_z^{n+1}|^2 + |e_\Gamma^{n+1}|^2] \left. \right\} \\
&\quad + \frac{1}{\gamma\Delta t} \left\{ 5|\mathcal{R}_{y_1}^n|^2 + \frac{5\Delta t^2}{4} |\mathcal{R}_{y_2}^n|^2 + 5|R_y^n|^2 \right\}.
\end{aligned} \tag{4.4}$$

• **Step 2:** Estimating the error  $e_z^n = Z_{t_n}^{t_n, X^n} - Z^n$ .

Subtracting the scheme (3.23) from the reference equation (3.16), we obtain

$$\begin{aligned}
\frac{\Delta t_n}{2} e_z^n &= \mathbb{E}_{t_n}^{X^n} \left[ (Y_{t_{n+1}}^{t_n, X^n} - Y^{n+1}) \Delta \tilde{W}_{t_{n+1}}^\top \right] \\
&\quad + \Delta t_n \mathbb{E}_{t_n}^{X^n} \left[ (f_{t_{n+1}}^{t_n, X^n} - f^{n+1}) \Delta \tilde{W}_{t_{n+1}}^\top \right] + R_z^n.
\end{aligned} \tag{4.5}$$

Substituting the identities

$$\begin{aligned}
\mathbb{E}_{t_n}^{X^n} \left[ (Y_{t_{n+1}}^{t_n, X^n} - Y^{n+1}) \Delta \tilde{W}_{t_{n+1}}^\top \right] &= \mathcal{R}_{z_1}^n + \mathbb{E}_{t_n}^{X^n} \left[ e_y^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top \right], \\
\mathbb{E}_{t_n}^{X^n} \left[ (f_{t_{n+1}}^{t_n, X^n} - f^{n+1}) \Delta \tilde{W}_{t_{n+1}}^\top \right] &= \mathcal{R}_{z_2}^n + \mathbb{E}_{t_n}^{X^n} \left[ e_f^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top \right],
\end{aligned}$$

into (4.5),  $|e_z^n|$  can be estimated by

$$\begin{aligned}
|e_z^n| &= \left| \frac{2}{\Delta t_n} \mathbb{E}_{t_n}^{X^n} \left[ e_y^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top \right] + 2 \mathbb{E}_{t_n}^{X^n} \left[ e_f^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top \right] \right. \\
&\quad \left. + \frac{2}{\Delta t_n} \mathcal{R}_{z_1}^n + 2\mathcal{R}_{z_2}^n + \frac{2}{\Delta t_n} R_z^n \right| \\
&\leq \frac{2}{\Delta t_n} \left| \mathbb{E}_{t_n}^{X^n} \left[ e_y^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top \right] \right| + 2 \left| \mathbb{E}_{t_n}^{X^n} \left[ e_f^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top \right] \right| \\
&\quad + \frac{2}{\Delta t_n} |\mathcal{R}_{z_1}^n| + 2|\mathcal{R}_{z_2}^n| + \frac{2}{\Delta t_n} |R_z^n|.
\end{aligned} \tag{4.6}$$

By Hölder's inequality and the inequality  $(a + b)^2 \leq (1 + \varepsilon)a^2 + (1 + \frac{1}{\varepsilon})b^2$  for any positive real

number  $\varepsilon$ , and from (4.6), we deduce

$$|e_z^n|^2 \leq (1 + \varepsilon) \left( \frac{2}{\Delta t_n} \right)^2 \left| \mathbb{E}_{t_n}^{X^n} \left[ e_y^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top \right] \right|^2 \quad (4.7)$$

$$+ \left( 1 + \frac{1}{\varepsilon} \right) \left\{ 2 \left| \mathbb{E}_{t_n}^{X^n} \left[ e_f^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top \right] \right| + \frac{2|\mathcal{R}_{z_1}^n|}{\Delta t_n} + 2|\mathcal{R}_{z_2}^n| + \frac{2|R_z^n|}{\Delta t_n} \right\}^2 \quad (4.8)$$

$$\leq (1 + \varepsilon) \left( \frac{2}{\Delta t_n} \right)^2 \left| \mathbb{E}_{t_n}^{X^n} \left[ e_y^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top \right] \right|^2 + 16 \left( 1 + \frac{1}{\varepsilon} \right) \left\{ \mathbb{E}_{t_n}^{X^n} \left[ |e_f^{n+1}|^2 \right] \mathbb{E}_{t_n}^{X^n} \left[ |\Delta \tilde{W}_{t_{n+1}}^\top|^2 \right] \right. \quad (4.9)$$

$$\left. + \left( \frac{1}{\Delta t_n} \right)^2 |\mathcal{R}_{z_1}^n|^2 + |\mathcal{R}_{z_2}^n|^2 + \left( \frac{1}{\Delta t_n} \right)^2 |R_z^n|^2 \right\}. \quad (4.10)$$

$$(4.11)$$

By the equality  $\mathbb{E}_{t_n}^{X^n} [|\Delta \tilde{W}_{t_{n+1}}|^2] = \Delta t_n$ , and the estimates of  $\mathbb{E}_{t_n}^{X^n} [|e_f^{n+1}|^2]$  and  $|\mathbb{E}_{t_n}^{X^n} [e_y^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top]|^2$ , i.e.,

$$\begin{aligned} \mathbb{E}_{t_n}^{X^n} [|e_f^{n+1}|^2] &\leq \mathbb{E}_{t_n}^{X^n} \left[ |L(|e_y^{n+1}| + |e_z^{n+1}| + |e_\Gamma^{n+1}|)|^2 \right] \\ &\leq 3L^2 \mathbb{E}_{t_n}^{X^n} [|e_y^{n+1}|^2 + |e_z^{n+1}|^2 + |e_\Gamma^{n+1}|^2], \end{aligned} \quad (4.12)$$

and

$$\begin{aligned} \left| \mathbb{E}_{t_n}^{X^n} [e_y^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top] \right|^2 &= \left| \mathbb{E}_{t_n}^{X^n} [(e_y^{n+1} - \mathbb{E}_{t_n}^{X^n} [e_y^{n+1}]) \Delta \tilde{W}_{t_{n+1}}^\top] \right|^2 \\ &\leq \mathbb{E}_{t_n}^{X^n} [|\Delta \tilde{W}_{t_{n+1}}^\top|^2] \mathbb{E}_{t_n}^{X^n} [(e_y^{n+1} - \mathbb{E}_{t_n}^{X^n} [e_y^{n+1}])^2] \\ &= \Delta t_n \left\{ \mathbb{E}_{t_n}^{X^n} [|e_y^{n+1}|^2] - \left| \mathbb{E}_{t_n}^{X^n} [e_y^{n+1}] \right|^2 \right\}, \end{aligned} \quad (4.13)$$

into (??), and dividing both sides of the resulting inequality by  $(1 + \varepsilon) \frac{8}{\Delta t}$ , we obtain an estimate of  $|e_z^n|^2$ , i.e.,

$$\begin{aligned} \frac{\Delta t}{8(1 + \varepsilon)} |e_z^n|^2 &\leq \frac{c_0}{2} \left\{ \mathbb{E}_{t_n}^{X^n} [|e_y^{n+1}|^2] - \left| \mathbb{E}_{t_n}^{X^n} [e_y^{n+1}] \right|^2 \right\} + \frac{6L^2}{\varepsilon} \Delta t^2 \mathbb{E}_{t_n}^{X^n} [|e_y^{n+1}|^2] \\ &\quad + |e_z^{n+1}|^2 + |e_\Gamma^{n+1}|^2 + \frac{2\Delta t}{\varepsilon} \left\{ \frac{|\mathcal{R}_{z_1}^n|^2}{(\Delta t_n)^2} + |\mathcal{R}_{z_2}^n|^2 + \frac{|R_z^n|^2}{(\Delta t_n)^2} \right\}. \end{aligned} \quad (4.14)$$

• **Step 3:** Estimating the error  $e_\Gamma^n = \Gamma_{t_n}^{t_n, X^n} - \Gamma^n$ .

Subtracting the scheme (3.24) and the reference equation (3.21), we have

$$\frac{1}{2} \Delta t_n e_\Gamma^n = \mathbb{E}_{t_n}^{X^n} \left[ (Y_{t_{n+1}}^{t_n, X^n} - Y^{n+1}) \Delta \tilde{\mu}_{t_{n+1}}^* \right] + \Delta t_n \mathbb{E}_{t_n}^{X^n} \left[ (f_{t_{n+1}}^{t_n, X^n} - f^{n+1}) \Delta \tilde{\mu}_{t_{n+1}}^* \right] + R_\Gamma^n. \quad (4.15)$$

Substituting

$$\begin{aligned} &\mathbb{E}_{t_n}^{X^n} \left[ (Y_{t_{n+1}}^{t_n, X^n} - Y^{n+1}) \Delta \tilde{\mu}_{t_{n+1}}^* \right] \\ &= \mathbb{E}_{t_n}^{X^n} \left[ (Y_{t_{n+1}}^{t_n, X^n} - Y_{t_{n+1}}^{t_{n+1}, X^{n+1}}) \Delta \tilde{\mu}_{t_{n+1}}^* \right] + \mathbb{E}_{t_n}^{X^n} \left[ e_y^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^* \right] \\ &= \mathcal{R}_{\Gamma_1}^n + \mathbb{E}_{t_n}^{X^n} \left[ e_y^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^* \right] \end{aligned}$$

and

$$\begin{aligned}
& \mathbb{E}_{t_n}^{X^n} \left[ (f_{t_{n+1}}^{t_n, X^n} - f^{n+1}) \Delta \tilde{\mu}_{t_{n+1}}^* \right] \\
&= \mathbb{E}_{t_n}^{X^n} \left[ (f_{t_{n+1}}^{t_n, X^n} - f_{t_{n+1}}^{t_{n+1}, X^{n+1}}) \Delta \tilde{\mu}_{t_{n+1}}^* \right] + \mathbb{E}_{t_n}^{X^n} \left[ e_f^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^* \right] \\
&= \mathcal{R}_{\Gamma_2}^n + \mathbb{E}_{t_n}^{X^n} \left[ e_f^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^* \right]
\end{aligned}$$

into (4.15), we obtain an expression of  $e_\Gamma^n$  as

$$\begin{aligned}
e_\Gamma^n &= \frac{2}{\Delta t_n} \mathbb{E}_{t_n}^{X^n} \left[ e_y^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^* \right] + 2 \mathbb{E}_{t_n}^{X^n} \left[ e_f^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^* \right] \\
&\quad + \frac{2}{\Delta t_n} \mathcal{R}_{\Gamma_1}^n + 2 \mathcal{R}_{\Gamma_2}^n + \frac{2}{\Delta t_n} R_\Gamma^n,
\end{aligned} \tag{4.16}$$

and consequently we obtain an upper bound of  $|e_\Gamma^n|$ , i.e.,

$$\begin{aligned}
|e_\Gamma^n| &\leq \frac{2}{\Delta t_n} \left| \mathbb{E}_{t_n}^{X^n} [e_y^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*] \right| + 2 \left| \mathbb{E}_{t_n}^{X^n} [e_f^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*] \right| \\
&\quad + \frac{2}{\Delta t_n} |\mathcal{R}_{\Gamma_1}^n| + 2 |\mathcal{R}_{\Gamma_2}^n| + \frac{2}{\Delta t_n} |R_\Gamma^n|.
\end{aligned} \tag{4.17}$$

By Hölder's inequality and the inequality  $(a+b)^2 \leq (1+\varepsilon)a^2 + (1+\frac{1}{\varepsilon})b^2$  for any positive real number  $\varepsilon$ , we obtain the following inequality from (4.17), i.e.,

$$\begin{aligned}
|e_\Gamma^n|^2 &\leq (1+\varepsilon) \left( \frac{2}{\Delta t_n} \right)^2 \left| \mathbb{E}_{t_n}^{X^n} [e_y^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*] \right|^2 \\
&\quad + \left( 1 + \frac{1}{\varepsilon} \right) \left\{ 2 \left| \mathbb{E}_{t_n}^{X^n} [e_f^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*] \right| + \frac{2 |\mathcal{R}_{\Gamma_1}^n|}{\Delta t_n} + 2 |\mathcal{R}_{\Gamma_2}^n| + \frac{2 |R_\Gamma^n|}{\Delta t_n} \right\}^2 \\
&\leq (1+\varepsilon) \left( \frac{2}{\Delta t_n} \right)^2 \left| \mathbb{E}_{t_n}^{X^n} [e_y^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*] \right|^2 \\
&\quad + 16 \left( 1 + \frac{1}{\varepsilon} \right) \left\{ \mathbb{E}_{t_n}^{X^n} [|e_f^{n+1}|^2] \mathbb{E}_{t_n}^{X^n} [|\Delta \tilde{\mu}_{t_{n+1}}^*|^2] \right. \\
&\quad \left. + \left( \frac{1}{\Delta t_n} \right)^2 |\mathcal{R}_{\Gamma_1}^n|^2 + |\mathcal{R}_{\Gamma_2}^n|^2 + \left( \frac{1}{\Delta t_n} \right)^2 |R_\Gamma^n|^2 \right\}.
\end{aligned} \tag{4.18}$$

By substituting the identity  $\mathbb{E}_{t_n}^{X^n} [|\Delta \tilde{\mu}_{t_{n+1}}^*|^2] = \Delta t_n \int_E \eta^2(e) \lambda(de)$ , the estimate of  $\mathbb{E}_{t_n}^{X^n} [|e_f^{n+1}|^2]$  given in (4.12), and the estimate  $|\mathbb{E}_{t_n}^{X^n} [e_y^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*]|^2$  into (4.18), and dividing both sides of the resulting inequality by  $\frac{8(1+\varepsilon) \int_E \rho^2(e) \lambda(de)}{\Delta t}$ , we deduce

$$\begin{aligned}
& \frac{\Delta t}{8(1+\varepsilon) \int_E \eta^2(e) \lambda(de)} |e_\Gamma^n|^2 \\
&\leq \frac{c_0}{2} \left\{ \mathbb{E}_{t_n}^{X^n} [|e_y^{n+1}|^2] - \left| \mathbb{E}_{t_n}^{X^n} [e_y^{n+1}] \right|^2 \right\} \\
&\quad + \frac{6L^2(\Delta t)^2}{\varepsilon} \mathbb{E}_{t_n}^{X^n} [|e_y^{n+1}|^2 + |e_z^{n+1}|^2 + |e_\Gamma^{n+1}|^2] \\
&\quad + \frac{2\Delta t}{\varepsilon \int_E \eta^2(e) \lambda(de)} \left[ \left( \frac{1}{\Delta t_n} \right)^2 |\mathcal{R}_{\Gamma_1}^n|^2 + |\mathcal{R}_{\Gamma_2}^n|^2 + \left( \frac{1}{\Delta t_n} \right)^2 |R_\Gamma^n|^2 \right].
\end{aligned} \tag{4.19}$$

• **Step 4:** Combining the estimates from Steps 1-3.

Now we add (4.4) multiplied by the constant  $c_0$ , (4.14), and (4.19) together to obtain the inequality

$$\begin{aligned}
& c_0 |e_y^n|^2 + \frac{\Delta t}{8(1+\varepsilon)} |e_z^n|^2 + \frac{\Delta t}{8(1+\varepsilon) \int_E \eta^2(e) \lambda(de)} |e_\Gamma^n|^2 \\
& \leq c_0 \left[ 1 + \left( \gamma + \frac{15L^2}{4\gamma} + \frac{15L^2\Delta t}{4} + \frac{12L^2\Delta t}{c_0\varepsilon} \right) \Delta t \right] \mathbb{E}_{t_n}^{X^n} [|e_y^{n+1}|^2] \\
& \quad + \left[ \frac{15c_0}{4\gamma} + \left( \frac{15c_0}{4} + \frac{12}{\varepsilon} \right) \Delta t \right] L^2 \Delta t \mathbb{E}_{t_n}^{X^n} [|e_z^{n+1}|^2] \\
& \quad + \left[ \frac{15c_0}{4\gamma} + \left( \frac{15c_0}{4} + \frac{12}{\varepsilon} \right) \Delta t \right] L^2 \Delta t \mathbb{E}_{t_n}^{X^n} [|e_\Gamma^{n+1}|^2] \\
& \quad + \left( \frac{15c_0}{4\gamma} + \frac{15c_0\Delta t}{4} \right) L^2 \Delta t (|e_y^n|^2 + |e_z^n|^2 + |e_\Gamma^n|^2) \\
& \quad + 5c_0 \left( 1 + \frac{1}{\gamma\Delta t} \right) \left\{ |\mathcal{R}_{y_1}^n|^2 + \frac{1}{4} \Delta t^2 |\mathcal{R}_{y_2}^n|^2 + |R_y^n|^2 \right\} \\
& \quad + \frac{2\Delta t}{\varepsilon} \left\{ \frac{1}{(\Delta t_n)^2} |\mathcal{R}_{z_1}^n|^2 + |\mathcal{R}_{z_2}^n|^2 + \frac{1}{(\Delta t_n)^2} |R_z^n|^2 \right\} \\
& \quad + \frac{2\Delta t}{\varepsilon \int_E \eta^2(e) \lambda(de)} \left\{ \frac{1}{(\Delta t_n)^2} |\mathcal{R}_{\Gamma_1}^n|^2 + |\mathcal{R}_{\Gamma_2}^n|^2 + \frac{1}{(\Delta t_n)^2} |R_\Gamma^n|^2 \right\}.
\end{aligned} \tag{4.20}$$

By taking expectation  $\mathbb{E}[\cdot]$  on both sides of (4.20), we deduce

$$\begin{aligned}
& c_0 (1 - C_1 \Delta t) \mathbb{E} [|e_y^n|^2] + C_3 \Delta t \mathbb{E} [|e_z^n|^2] + C_6 \Delta t \mathbb{E} [|e_\Gamma^n|^2] \\
& \leq c_0 (1 + C_2 \Delta t) \mathbb{E} [|e_y^{n+1}|^2] + C_4 \Delta t \mathbb{E} [|e_z^{n+1}|^2] + C_4 \Delta t \mathbb{E} [|e_\Gamma^{n+1}|^2] \\
& \quad + \frac{C_5}{\Delta t} \mathbb{E} \left[ |\mathcal{R}_{y_1}^n|^2 + \frac{1}{4} \Delta t^2 |\mathcal{R}_{y_2}^n|^2 + |R_y^n|^2 \right] \\
& \quad + \frac{2\Delta t}{\varepsilon} \mathbb{E} \left[ \left( \frac{1}{\Delta t_n} \right)^2 |\mathcal{R}_{z_1}^n|^2 + |\mathcal{R}_{z_2}^n|^2 + \left( \frac{1}{\Delta t_n} \right)^2 |R_z^n|^2 \right] \\
& \quad + \frac{2\Delta t}{\varepsilon \int_E \eta^2(e) \lambda(de)} \mathbb{E} \left[ \frac{1}{(\Delta t_n)^2} |\mathcal{R}_{\Gamma_1}^n|^2 + |\mathcal{R}_{\Gamma_2}^n|^2 + \frac{1}{(\Delta t_n)^2} |R_\Gamma^n|^2 \right],
\end{aligned} \tag{4.21}$$

where the constants  $C_1, C_2, C_3, C_4, C_5, C_6$  are defined by

$$\begin{aligned}
C_1 &= \left( \frac{15}{4\gamma} + \frac{15\Delta t}{4} \right) L^2, \quad C_2 = \left( \gamma + \frac{15L^2}{4\gamma} + \frac{15L^2\Delta t}{4} + \frac{12L^2\Delta t}{c_0\varepsilon} \right), \\
C_3 &= \frac{1}{8(1+\varepsilon)} - \left( \frac{15c_0}{4\gamma} + \frac{15c_0\Delta t}{4} \right) L^2, \quad C_4 = \left[ \frac{15c_0}{4\gamma} + \left( \frac{15c_0}{4} + \frac{12}{\varepsilon} \right) \Delta t \right] L^2, \\
C_5 &= 5c_0 \frac{1 + \gamma\Delta t}{\gamma}, \quad C_6 = \frac{1}{8(1+\varepsilon) \int_E \eta^2(e) \lambda(de)} - \left( \frac{15c_0}{4\gamma} + \frac{15c_0\Delta t}{4} \right) L^2.
\end{aligned} \tag{4.22}$$

Now we set  $\varepsilon = 1$ ,  $\gamma$  large enough and  $\Delta t$  sufficiently small, such that if  $0 < \Delta t \leq \Delta t_0$  then  $C_1 \leq C$ ,  $C_2 \leq C$ ,  $C_5 \leq C$ ,  $1 - C\Delta t > 0$ , and  $C_3 - C_4 > C^* > 0$ ,  $C_6 - C_4 > C^* > 0$  where  $C$  and  $C^*$  are two positive constants depending on  $c_0$  and  $L$ . Then for  $0 < \Delta t \leq \Delta t_0$ , we deduce

from (4.21)

$$\begin{aligned}
& c_0(1 - C\Delta t)\mathbb{E}[|e_y^n|^2] + C_3\Delta t\mathbb{E}[|e_z^n|^2] + C_6\Delta t\mathbb{E}[|e_\Gamma^n|^2] \\
& \leq c_0(1 + C\Delta t)\mathbb{E}[|e_y^{n+1}|^2] + C_4\Delta t\mathbb{E}[|e_z^{n+1}|^2] + C_4\Delta t\mathbb{E}[|e_\Gamma^{n+1}|^2] \\
& \quad + C\mathbb{E}\left[\frac{1}{\Delta t}|\mathcal{R}_{y_1}^n|^2 + \Delta t|\mathcal{R}_{y_2}^n|^2 + \frac{1}{\Delta t}|R_y^n|^2\right] \\
& \quad + C\mathbb{E}\left[\frac{1}{\Delta t}|\mathcal{R}_{z_1}^n|^2 + \Delta t|\mathcal{R}_{z_2}^n|^2 + \frac{1}{\Delta t}|R_z^n|^2\right] \\
& \quad + C\mathbb{E}\left[\frac{1}{\Delta t}|\mathcal{R}_{\Gamma_1}^n|^2 + \Delta t|\mathcal{R}_{\Gamma_2}^n|^2 + \frac{1}{\Delta t}|R_\Gamma^n|^2\right].
\end{aligned}$$

Dividing both sides of the upper inequality by  $(1 - C\Delta t)$ , we easily get

$$\begin{aligned}
& c_0\mathbb{E}[|e_y^n|^2] + C_3\Delta t\mathbb{E}[|e_z^n|^2] + C_6\Delta t\mathbb{E}[|e_\Gamma^n|^2] \\
& \leq \frac{1 + C\Delta t}{1 - C\Delta t} \left( c_0\mathbb{E}[|e_y^{n+1}|^2] + C_4\Delta t\mathbb{E}[|e_z^{n+1}|^2] + C_4\Delta t\mathbb{E}[|e_\Gamma^{n+1}|^2] \right) \\
& \quad + \frac{C}{1 - C\Delta t} \mathbb{E}\left[\frac{1}{\Delta t}|\mathcal{R}_{y_1}^n|^2 + \Delta t|\mathcal{R}_{y_2}^n|^2 + \frac{1}{\Delta t}|R_y^n|^2\right] \\
& \quad + \frac{C}{1 - C\Delta t} \mathbb{E}\left[\frac{1}{\Delta t}|\mathcal{R}_{z_1}^n|^2 + \Delta t|\mathcal{R}_{z_2}^n|^2 + \frac{1}{\Delta t}|R_z^n|^2\right] \\
& \quad + \frac{C}{1 - C\Delta t} \mathbb{E}\left[\frac{1}{\Delta t}|\mathcal{R}_{\Gamma_1}^n|^2 + \Delta t|\mathcal{R}_{\Gamma_2}^n|^2 + \frac{1}{\Delta t}|R_\Gamma^n|^2\right].
\end{aligned} \tag{4.23}$$

From the inequality (4.23), by recursively inserting  $e_y^i$ ,  $i = n + 1, \dots, N - 1$ , we deduce

$$\begin{aligned}
& c_0\mathbb{E}[|e_y^n|^2] + C_3\Delta t \sum_{i=n}^{N-1} \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{i-n} \mathbb{E}[|e_z^i|^2] + C_6\Delta t \sum_{i=n}^{N-1} \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{i-n} \mathbb{E}[|e_\Gamma^i|^2] \\
& \leq \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{N-n} c_0\mathbb{E}[|e_y^N|^2] + C_4\Delta t \sum_{i=n+1}^N \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{i-n} \mathbb{E}[|e_z^i|^2] \\
& \quad + C_4\Delta t \sum_{i=n+1}^N \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{i-n} \mathbb{E}[|e_\Gamma^i|^2] \\
& \quad + \sum_{i=n}^{N-1} \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{i-n} \frac{C}{1 - C\Delta t} \mathbb{E}\left[\frac{1}{\Delta t}|\mathcal{R}_{y_1}^i|^2 + \Delta t|\mathcal{R}_{y_2}^i|^2 + \frac{1}{\Delta t}|R_y^i|^2\right] \\
& \quad + \sum_{i=n}^{N-1} \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{i-n} \frac{C}{1 - C\Delta t} \mathbb{E}\left[\frac{1}{\Delta t}|\mathcal{R}_{z_1}^i|^2 + \Delta t|\mathcal{R}_{z_2}^i|^2 + \frac{1}{\Delta t}|R_z^i|^2\right] \\
& \quad + \sum_{i=n}^{N-1} \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{i-n} \frac{C}{1 - C\Delta t} \mathbb{E}\left[\frac{1}{\Delta t}|\mathcal{R}_{\Gamma_1}^i|^2 + \Delta t|\mathcal{R}_{\Gamma_2}^i|^2 + \frac{1}{\Delta t}|R_\Gamma^i|^2\right],
\end{aligned}$$

which immediately leads to

$$c_0\mathbb{E}[|e_y^n|^2] + C^*\Delta t \sum_{i=n}^{N-1} \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{i-n} \mathbb{E}[|e_z^i|^2] + C^*\Delta t \sum_{i=n}^{N-1} \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{i-n} \mathbb{E}[|e_\Gamma^i|^2]$$



$$\begin{aligned}
&\leq \left(\frac{1+C\Delta t}{1-C\Delta t}\right)^{N-n} c_0 \mathbb{E}[|e_y^N|^2] + C_4 \Delta t \left(\frac{1+C\Delta t}{1-C\Delta t}\right)^{N-n} \mathbb{E}[|e_z^N|^2] \\
&\quad + C_4 \Delta t \left(\frac{1+C\Delta t}{1-C\Delta t}\right)^{N-n} \mathbb{E}[|e_\Gamma^N|^2] \\
&\quad + \sum_{i=n}^{N-1} \left(\frac{1+C\Delta t}{1-C\Delta t}\right)^{i-n} \frac{C}{1-C\Delta t} \mathbb{E}\left[\frac{1}{\Delta t} |\mathcal{R}_{y_1}^i|^2 + \Delta t |\mathcal{R}_{y_2}^i|^2 + \frac{1}{\Delta t} |R_y^i|^2\right] \\
&\quad + \sum_{i=n}^{N-1} \left(\frac{1+C\Delta t}{1-C\Delta t}\right)^{i-n} \frac{C}{1-C\Delta t} \mathbb{E}\left[\frac{1}{\Delta t} |\mathcal{R}_{z_1}^i|^2 + \Delta t |\mathcal{R}_{z_2}^i|^2 + \frac{1}{\Delta t} |R_z^i|^2\right] \\
&\quad + \sum_{i=n}^{N-1} \left(\frac{1+C\Delta t}{1-C\Delta t}\right)^{i-n} \frac{C}{1-C\Delta t} \mathbb{E}\left[\frac{1}{\Delta t} |\mathcal{R}_{\Gamma_1}^i|^2 + \Delta t |\mathcal{R}_{\Gamma_2}^i|^2 + \frac{1}{\Delta t} |R_\Gamma^i|^2\right].
\end{aligned}$$

The proof is completed.

**Remark 1.** It is worth to note that Theorem 4.1 also implies that Scheme 1 is stable with respect to the terminal condition, that is, for any  $\varepsilon > 0$ , there exists a positive number  $\delta > 0$ , such that, if  $\mathbb{E}[|\bar{Y}^N - Y^N|^2] < \delta$ ,  $\mathbb{E}[|\bar{Z}^N - Z^N|^2] < \delta$  and  $\mathbb{E}[|\bar{\Gamma}^N - \Gamma^N|^2] < \delta$ , where  $(\bar{Y}^N, \bar{Z}^N, \bar{\Gamma}^N)$  and  $(Y^N, Z^N, \Gamma^N)$  are two different terminal conditions, then for  $0 \leq n \leq N-1$ , it holds that

$$\mathbb{E}[|\bar{Y}^n - Y^n|^2] + \Delta t \sum_{i=n}^N \mathbb{E}[|\bar{Z}^i - Z^i|^2] + \Delta t \sum_{i=n}^N \mathbb{E}[|\bar{\Gamma}^i - \Gamma^i|^2] < \varepsilon.$$

The next task is to estimate all the residual terms in (4.3), and the main technique used is the Itô-Taylor expansion [25]. Under some reasonable regularity conditions on the data  $b$ ,  $\sigma$ ,  $c$ ,  $f$  and  $\varphi$  in the FBSDE, we now derive estimates for the local truncation errors  $R_y^n$ ,  $R_z^n$  and  $R_\Gamma^n$  defined in (3.11), (3.16) and (3.21), respectively. To proceed, we need the following standard assumption.

**Assumption 4.1.** Under the condition that  $X_0$  is  $\mathcal{A}_0$ -measurable as well as  $\mathbb{E}[|X_0|^2] < \infty$ , we assume that  $b$ ,  $\sigma$  and  $c$  are jointly  $L^2$ -measurable in  $(t, x) \in [0, T] \times \mathbb{R}^q$ , and there exist real constants  $L > 0$  and  $K > 0$  such that

$$\begin{aligned}
|b(t, x) - b(t, x')| &\leq L|x - x'|, \quad |\sigma(t, x) - \sigma(t, x')| \leq L|x - x'|, \\
\int_E |c(t, x, e) - c(t, x', e)|^2 \lambda(de) &\leq L|x - x'|^2,
\end{aligned} \tag{4.24}$$

and

$$\begin{aligned}
|b(t, x)|^2 &\leq K(1 + |x|^2), \quad |\sigma(t, x)|^2 \leq K(1 + |x|^2), \\
\int_E |c(t, x, e)|^2 \lambda(de) &\leq K(1 + |x|^2),
\end{aligned} \tag{4.25}$$

for all  $t \in [0, T]$  and  $x, x' \in \mathbb{R}^q$ .

Under Assumption 4.1, if  $\mathbb{E}[|X_0|^{2m}] < \infty$  for some integer  $m \geq 1$ , the solution of the forward SDE in (2.1) has the estimate

$$\mathbb{E}_{t_n}^{X^n} [ |X_s^{t_n, X^n}|^{2m} ] \leq \left(1 + \mathbb{E}_{t_n}^{X^n} [ |X^n|^{2m} ]\right) e^{C(s-t_n)}, \tag{4.26}$$

where  $s \in [t_n, T]$  and  $C$  is a positive constant depending only on the constants  $K, L$  and  $m$ .

For the sake of presentation simplicity, in the following lemmas and theorems, we only consider the one-dimensional case ( $q = d = 1$ ), but our results can be extended to multidimensional cases without any essential difficulty. To proceed, we define three partial integro-differential operators:

$$\begin{aligned} L^0 v(t, x) &:= \frac{\partial v}{\partial t}(t, x) + b(t, x) \frac{\partial v}{\partial x}(t, x) + \frac{1}{2} \sigma^2(t, x) \frac{\partial^2 v}{\partial x^2}(t, x) \\ &\quad + \int_E \left[ v(t, x + c(t, x, e)) - v(t, x) - \frac{\partial v}{\partial x}(t, x) c(t, x, e) \right] \lambda(de), \end{aligned} \quad (4.27)$$

$$L^1 v(t, x) := \sigma(t, x) \frac{\partial v}{\partial x}(t, x),$$

$$L^{-1} v(t, x) := v(t, x + c(t, x, e)) - v(t, x);$$

and introduce the following notation:

$$\begin{aligned} &\mathcal{C}_b^{(k_1, \dots, k_J)}(D_1 \times \dots \times D_J) \\ &:= \left\{ \phi : \prod_{j=1}^J D_j \rightarrow \mathbb{R} \mid \frac{\partial^{\alpha_1} \dots \partial^{\alpha_J} \phi}{\partial^{\alpha_1} x_1 \dots \partial^{\alpha_J} x_J} \text{ is bounded and continuous} \right. \\ &\quad \left. \text{for } \alpha_j \leq k_j, j = 1, \dots, J, \text{ where } \vec{\alpha} := (\alpha_1, \dots, \alpha_J) \in \mathbb{N}^J \right\}, \end{aligned}$$

where  $J \in \mathbb{N}^+$ ,  $D_1 \times \dots \times D_J \subset \mathbb{R}^J$ .

Now we give the estimates of  $R_y^n, R_{z,1}^n, R_{z,2}^n, R_{\Gamma,1}^n$  and  $R_{\Gamma,2}^n$  in the following Lemma 1.

**Lemma 1.** *Under Assumption 4.1, if the data of the FBSDEs in (2.1) satisfy the following regularity conditions:  $f(t, x, y, z, \gamma) \in \mathcal{C}_b^{(2,4,4,4,4)}([0, T] \times \mathbb{R}^4)$ ,  $b(t, x) \in \mathcal{C}_b^{(2,4)}([0, T] \times \mathbb{R})$ ,  $\sigma(t, x) \in \mathcal{C}_b^{(2,4)}([0, T] \times \mathbb{R})$ ,  $\varphi(x) \in \mathcal{C}_b^{6+\alpha}(\mathbb{R})$  with  $\alpha \in (0, 1)$ , and  $c(t, x, e) \in \mathcal{C}_b^{(2,4,\infty)}([0, T] \times \mathbb{R}^2)$ , then for sufficiently small  $\Delta t = \max_n \Delta t_n$ , we have the estimates*

$$\mathbb{E} [|R_y^n|^2] \leq C (1 + \mathbb{E} [|X^n|^8]) (\Delta t)^6,$$

$$\mathbb{E} [|R_{z,1}^n|^2] \leq C (1 + \mathbb{E} [|X^n|^8]) (\Delta t_n)^6, \quad \mathbb{E} [|R_{z,2}^n|^2] \leq C (1 + \mathbb{E} [|X^n|^8]) (\Delta t_n)^6,$$

$$\mathbb{E} [|R_{\Gamma,1}^n|^2] \leq C (1 + \mathbb{E} [|X^n|^8]) (\Delta t_n)^6, \quad \mathbb{E} [|R_{\Gamma,2}^n|^2] \leq C (1 + \mathbb{E} [|X^n|^8]) (\Delta t_n)^6,$$

where  $R_y^n, R_{z,1}^n, R_{z,2}^n, R_{\Gamma,1}^n$  and  $R_{\Gamma,2}^n$  are defined in (3.10), (3.14), (3.15), (3.19) and (3.20), respectively, and  $C$  is a positive constant depending only on  $T, K$  and upper bounds of the derivatives of  $b, \sigma, c, f$  and  $\varphi$ .

*Proof.* Based on the relation shown in (2.5) between the solution  $(Y_t, Z_t, \Gamma_t)$  of the BSDE in (2.4) and the solution  $u(t, x)$  of the PIDE in (2.2), it is easy to prove that  $u(t, x) \in \mathcal{C}_b^{(2,4)}([0, T] \times \mathbb{R}^q)$  under the regularity conditions given in [1] on  $f, b, \sigma, c$  and  $\varphi$ . Then, for  $t \leq s \leq T$ , the function  $F = F(t, x)$  defined by

$$F(s, x) := f(s, x, u(s, x), \nabla u(s, x) \sigma(s, x), \Gamma(s, x)), \quad (4.28)$$

is in the space  $\mathcal{C}_b^{(2,4)}([0, T] \times \mathbb{R}^q)$ . Setting  $x = X_s^{t_n, X^n}$  in (4.28) and applying the Itô-Taylor

expansion to  $F(s, X_s^{t_n, X^n})$ , we obtain

$$\begin{aligned} F(s, X_s^{t_n, X^n}) &= F(t_n, X^n) + \int_{t_n}^s L^0 F(r, X_r^{t_n, X^n}) dr + \int_{t_n}^s L^1 F(r, X_r^{t_n, X^n}) dW_r \\ &\quad + \int_{t_n}^s \int_E L^{-1} F(r, X_{r-}^{t_n, X^n}) \tilde{\mu}(de, dr), \end{aligned} \quad (4.29)$$

where the operators  $L^0$ ,  $L^1$  and  $L^{-1}$  are defined in (4.27).

Taking the conditional expectation  $\mathbb{E}_{t_n}^{X^n}[\cdot]$  on both sides, we have that

$$\mathbb{E}_{t_n}^{X^n} \left[ \int_{t_n}^s L^1 F(r, X_r^{t_n, X^n}) dW_r \right] = 0 \quad \text{and} \quad \mathbb{E}_{t_n}^{X^n} \left[ \int_{t_n}^s \int_E L^{-1} F(r, X_{r-}^{t_n, X^n}) \tilde{\mu}(de, dr) \right] = 0.$$

Then following the same arguments used in the proof of Lemmas 4.2-4.4 in [36], we obtain the estimates of the lemma. The proof is completed.

From Lemma 1 and the definitions of  $R_z^n$  in (3.16) and  $R_\Gamma^n$  in (3.21), we easily get the estimates of  $R_z^n$  and  $R_\Gamma^n$ , stated in the following lemma.

**Lemma 2.** *Under the conditions of Lemma 1, for sufficiently small time step size  $\Delta t = \max_n \Delta t_n$ , we have that*

$$\mathbb{E} [|R_z^n|^2] \leq C (1 + \mathbb{E} [|X^n|^8]) (\Delta t)^6 \quad \text{for } 0 \leq n \leq N-1,$$

$$\mathbb{E} [|R_\Gamma^n|^2] \leq C (1 + \mathbb{E} [|X^n|^8]) (\Delta t)^6 \quad \text{for } 0 \leq n \leq N-1,$$

where  $R_z^n$ ,  $R_\Gamma^n$  are defined in (3.16), (3.21),  $C$  is a positive constant depending on  $T$ ,  $K$  and the upper bounds of the derivatives of  $b$ ,  $\sigma$ ,  $c$ ,  $f$  and  $\varphi$ .

Now combining Theorem 4.1, Lemma 1 and Lemma 2 as well as the estimates given in (3.5) and (3.6), we obtain the convergence rate of Scheme 1 in the following theorem.

**Theorem 4.2.** *Under Lemma 1 and Lemma 2, if (3.5) and (3.6) hold for the scheme (3.4) for the forward SDE, then, for sufficiently small time step size  $\Delta t = \max_n \Delta t_n$ , the errors  $e_y^n$ ,  $e_z^n$  and  $e_\Gamma^n$  in (4.1) for  $n = 0, \dots, N-1$  can be bounded by*

$$\begin{aligned} &\mathbb{E} [|e_y^n|^2] + \Delta t \sum_{i=n}^{N-1} \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{i-n} \mathbb{E} [|e_z^i|^2 + |e_\Gamma^i|^2] \\ &\leq C_1 (\mathbb{E} [|e_y^N|^2] + \Delta t \mathbb{E} [|e_z^N|^2 + |e_\Gamma^N|^2]) + C_2 \left( (\Delta t)^{2\alpha} + (\Delta t)^{2\beta} + (\Delta t)^{2\gamma} + (\Delta t)^4 \right), \end{aligned}$$

where  $\alpha$ ,  $\beta$ ,  $\gamma$  are defined in (3.6),  $C > 0$  depends on  $c_0$  and  $L$ ,  $C_1 > 0$  depends on  $c_0$ ,  $T$  and  $L$ ,  $C_2 > 0$  depends on  $c_0$ ,  $T$ ,  $L$ ,  $K$ ,  $X_0$  and the upper bounds of the derivatives of  $b$ ,  $\sigma$ ,  $c$ ,  $f$  and  $\varphi$ .

*Proof.* From the definitions of  $\mathcal{R}_{y_1}^i$ ,  $\mathcal{R}_{y_2}^i$ ,  $\mathcal{R}_{z_1}^i$ ,  $\mathcal{R}_{z_2}^i$ ,  $\mathcal{R}_{\Gamma_1}^i$  and  $\mathcal{R}_{\Gamma_2}^i$  in Theorem 4.1, under the

conditions of the theorem, we easily get the estimates

$$\begin{aligned}
& \mathbb{E}[|X^i|^2] \leq C(1 + \mathbb{E}[|X_0|^2]), \\
& \mathbb{E}[|\mathcal{R}_{y_1}^i|^2] \leq C(1 + \mathbb{E}[|X^i|^{4r_1}])(\Delta t)^{2\beta+2} \leq C(1 + \mathbb{E}[|X_0|^{4r_1}])(\Delta t)^{2\beta+2}, \\
& \mathbb{E}[|\mathcal{R}_{y_2}^i|^2] \leq C(1 + \mathbb{E}[|X^i|^{4r_1}])(\Delta t)^{2\beta+2} \leq C(1 + \mathbb{E}[|X_0|^{4r_1}])(\Delta t)^{2\beta+2}, \\
& \mathbb{E}[|\mathcal{R}_{z_1}^i|^2] \leq C(1 + \mathbb{E}[|X^i|^{4r_2}])(\Delta t)^{2\gamma+2} \leq C(1 + \mathbb{E}[|X_0|^{4r_2}])(\Delta t)^{2\gamma+2}, \\
& \mathbb{E}[|\mathcal{R}_{z_2}^i|^2] \leq C(1 + \mathbb{E}[|X^i|^{4r_2}])(\Delta t)^{2\gamma+2} \leq C(1 + \mathbb{E}[|X_0|^{4r_2}])(\Delta t)^{2\gamma+2}, \\
& \mathbb{E}[|\mathcal{R}_{\Gamma_1}^i|^2] \leq C(1 + \mathbb{E}[|X^i|^{4r_3}])(\Delta t)^{2\alpha+2} \leq C(1 + \mathbb{E}[|X_0|^{4r_3}])(\Delta t)^{2\alpha+2}, \\
& \mathbb{E}[|\mathcal{R}_{\Gamma_2}^i|^2] \leq C(1 + \mathbb{E}[|X^i|^{4r_3}])(\Delta t)^{2\alpha+2} \leq C(1 + \mathbb{E}[|X_0|^{4r_3}])(\Delta t)^{2\alpha+2}
\end{aligned} \tag{4.30}$$

for  $i = 0, 1, \dots, N-1$ . By Lemmas 1, 2 and 2, and inequality (3.5), for  $0 \leq i \leq N-1$ , we have

$$\begin{aligned}
& \mathbb{E}[|R_y^i|^2] \leq C(1 + \mathbb{E}[|X_0|^8])(\Delta t)^6, \quad \mathbb{E}[|R_z^i|^2] \leq C(1 + \mathbb{E}[|X_0|^8])(\Delta t)^6, \\
& \mathbb{E}[|R_\Gamma^i|^2] \leq C(1 + \mathbb{E}[|X_0|^8])(\Delta t)^6.
\end{aligned} \tag{4.31}$$

By (4.30) and (4.31), we deduce

$$\begin{aligned}
& \sum_{i=n}^{N-1} \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{i-n} \frac{C\mathbb{E}[|\mathcal{R}_{y_1}^i|^2 + (\Delta t)^2|\mathcal{R}_{y_2}^i|^2 + |R_y^i|^2]}{\Delta t(1 - C\Delta t)} \\
& \leq C(1 + \mathbb{E}[|X_0|^{4r_1}] + \mathbb{E}[|X_0|^8])(\Delta t)^{2\beta} + (\Delta t)^4,
\end{aligned} \tag{4.32}$$

$$\begin{aligned}
& \sum_{i=n}^{N-1} \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{i-n} \frac{C\mathbb{E}[|\mathcal{R}_{z_1}^i|^2 + (\Delta t)^2|\mathcal{R}_{z_2}^i|^2 + |R_z^i|^2]}{\Delta t(1 - C\Delta t)} \\
& \leq C(1 + \mathbb{E}[|X_0|^{4r_2}] + \mathbb{E}[|X_0|^8])(\Delta t)^{2\gamma} + (\Delta t)^4,
\end{aligned} \tag{4.33}$$

and

$$\begin{aligned}
& \sum_{i=n}^{N-1} \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{i-n} \frac{C\mathbb{E}[|\mathcal{R}_{\Gamma_1}^i|^2 + (\Delta t)^2|\mathcal{R}_{\Gamma_2}^i|^2 + |R_\Gamma^i|^2]}{\Delta t(1 - C\Delta t)} \\
& \leq C(1 + \mathbb{E}[|X_0|^{4r_3}] + \mathbb{E}[|X_0|^8])(\Delta t)^{2\alpha} + (\Delta t)^4.
\end{aligned} \tag{4.34}$$

By applying (4.32), (4.33) and (4.34) to Theorem 4.1, we complete the proof.

## 5. The fully discrete scheme

In this section, we will develop a fully discrete scheme based on Scheme 1 by assuming that the jump process of  $X_t$  in (2.1) has *finite activity*. This means the Poisson random measure  $\tilde{\mu}(de, dt)$  can be represented by

$$\tilde{\mu}(de, dt) = \mu(de, dt) - \lambda \rho(e) de dt, \tag{5.1}$$

where  $0 < \lambda < \infty$  is the jump intensity and  $\rho(e)de$  is the probability measure of each jump size satisfying  $\int_E \rho(e)de = 1$ . For jump processes with infinite activities, i.e.,  $\lambda = \infty$ , substantial efforts are needed to construct new spatial discretization approaches, which is out of scope of this paper and will be considered in our future works.

To proceed, we first introduce a partition the  $q$ -dimensional Euclidean space  $\mathbb{R}^q$  by  $\mathcal{S} =$

$\mathcal{S}^1 \times \mathcal{S}^2 \times \cdots \times \mathcal{S}^q$ , where  $\mathcal{S}^k$  for  $k = 1, \dots, q$  is a partition of the one-dimensional space  $\mathbb{R}$ , i.e.,

$$\mathcal{S}^k = \left\{ x_i^k \mid x_i^k \in \mathbb{R}, i \in \mathbb{Z}, x_i^k < x_{i+1}^k, \lim_{i \rightarrow +\infty} x_i^k = +\infty, \lim_{i \rightarrow -\infty} x_i^k = -\infty \right\},$$

where  $\Delta x_k := \max_{i \in \mathbb{Z}} \{|x_i^k - x_{i-1}^k|\}$  and  $\Delta x = \max_{1 \leq k \leq q} \Delta x_k$ . For each multi-index  $\mathbf{i} = (i_1, i_2, \dots, i_q) \in \mathbb{Z}^q$ , the corresponding grid point in  $\mathcal{S}$  is denoted by  $x_{\mathbf{i}} = (x_{i_1}^1, \dots, x_{i_q}^q)$ .

Recalling (2.5) and (3.26), we can see that, if the terminal condition  $\xi$  is a function of  $X_T$ ,  $(Y_t^{t,x}, Z_t^{t,x}, \Gamma_t^{t,x})$  can be treated as functions of  $t$  and  $X_t^{t,x} = x$  for  $0 \leq t \leq T$ . Analogously, the semi-discrete solution  $(Y^n, Z^n, \Gamma^n)$  can be treated as functions of  $X^n$ . Thus, in this section, we also write  $(Y^n, Z^n, \Gamma^n)$  as functions of  $x \in \mathbb{R}^q$ , i.e.,

$$Y^n(x) := \mathbb{E}[Y^n | X^n = x], \quad Z^n(x) := \mathbb{E}[Z^n | X^n = x], \quad \Gamma^n(x) := \mathbb{E}[\Gamma^n | X^n = x].$$

Our objective is to approximate the exact solution  $(Y_{t_n}^{t_n, x_{\mathbf{i}}}, Z_{t_n}^{t_n, x_{\mathbf{i}}}, \Gamma_{t_n}^{t_n, x_{\mathbf{i}}})$  by constructing  $(Y_{\mathbf{i}}^n, Z_{\mathbf{i}}^n, \Gamma_{\mathbf{i}}^n)$ , such that

$$Y_{\mathbf{i}}^n \approx Y^n(x_{\mathbf{i}}) \approx Y_{t_n}^{t_n, x_{\mathbf{i}}}, \quad Z_{\mathbf{i}}^n \approx Z^n(x_{\mathbf{i}}) \approx Z_{t_n}^{t_n, x_{\mathbf{i}}}, \quad \Gamma_{\mathbf{i}}^n \approx \Gamma^n(x_{\mathbf{i}}) \approx \Gamma_{t_n}^{t_n, x_{\mathbf{i}}},$$

for  $n = 0, \dots, N-1$  and  $\mathbf{i} \in \mathbb{Z}^q$ .

To this end, it is critical to develop effective quadrature rules for approximating the conditional mathematical expectations  $\mathbb{E}_{t_n}^{X^n}[\cdot]$  in (3.23)-(3.25). For instance, at each time-space point  $(t_n, x_{\mathbf{i}}) \in \mathcal{T} \times \mathcal{S}$ , approximating  $Y^n(x_{\mathbf{i}})$  using (3.25) requires quadrature rules for  $\mathbb{E}_{t_n}^{x_{\mathbf{i}}}[Y^{n+1}]$  and  $\mathbb{E}_{t_n}^{x_{\mathbf{i}}}[f^{n+1}]$ . In what follows, we take  $\mathbb{E}_{t_n}^{x_{\mathbf{i}}}[Y^{n+1}]$  as an example to propose our new quadrature rule. Slight modifications are needed for approximating  $\mathbb{E}_{t_n}^{x_{\mathbf{i}}}[Y^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top]$  and  $\mathbb{E}_{t_n}^{x_{\mathbf{i}}}[Y^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*]$ ; all the proposed quadrature rules can be directly used to estimate the expectations of  $f^{n+1}$ ,  $f^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top$  and  $f^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*$ .

It is observed that  $\mathbb{E}_{t_n}^{x_{\mathbf{i}}}[Y^{n+1}]$  is defined with respect to the probability measure of the incremental stochastic process  $\Delta X^{n+1} := X^{n+1} - x_{\mathbf{i}} = \Phi(t_n, t_{n+1}, x_{\mathbf{i}}, I_{\mathcal{J} \in \mathcal{A}_\beta})$  starting from  $(t_n, x_{\mathbf{i}})$ , where  $\Phi$  is determined by the selected scheme for the forward SDE. In this section, for the sake of simplicity, we choose the forward Euler method for the scheme in (3.22), i.e.,

$$X^{n+1} = x_{\mathbf{i}} + b(t_n, x_{\mathbf{i}}) \Delta t_n + \sigma(t_n, x_{\mathbf{i}}) \Delta W_{t_{n+1}} + \sum_{k=N_{t_n}+1}^{N_{t_{n+1}}} c(t_n, x_{\mathbf{i}}, e_k), \quad (5.2)$$

where  $N_t$  for  $t \in [t_n, t_{n+1}]$  is the underlying Poisson process. Since (5.2) only achieves first-order convergence in the weak sense, the overall convergence of Scheme 1 will be of first order. High-order schemes for the forward SDE [25], such as order-2.0 weak Taylor scheme, can also be used, but the corresponding quadrature rules for approximating  $\mathbb{E}_{t_n}^{x_{\mathbf{i}}}[\cdot]$  will be dramatically different from the case of using (5.2). Since the jump intensity  $\lambda$  in (5.1) is finite, the number of jumps of  $X_t$  within  $(t_n, t_{n+1}]$  follows a compensated Poisson distribution  $N_{t_{n+1}} - N_{t_n} - \lambda \Delta t_n$ , where the size of each jump, i.e.,  $c(t_n, x_{\mathbf{i}}, e)$ , follows the distribution  $\rho(e)de$ . Next, we observe that

$$\Delta W_{t_{n+1}}^i = \sqrt{\Delta t_n} \xi^i \quad \text{for } i = 1, \dots, d, \quad (5.3)$$

where  $\xi^i$  follows the standard normal distribution  $N(0, 1)$ . Hereafter, we denote by  $\varrho(\xi^i)$  the probability density function of  $\xi^i$ , and by  $\varrho^d(\xi)$  the joint probability density function of  $\xi = (\xi^1, \dots, \xi^d)^\top$ .

Now, we can write out the expression of  $\mathbb{E}_{t_n}^{x_i}[Y^{n+1}]$  as

$$\begin{aligned}
& \mathbb{E}_{t_n}^{x_i}[Y^{n+1}] \\
&= \sum_{m=0}^{\infty} \mathbb{P}\{N_{t_{n+1}} - N_{t_n} = m\} \mathbb{E}\left[Y^{n+1}\left(x_i + b(t_n, x_i)\Delta t_n + \sigma(t_n, x_i)\sqrt{\Delta t_n}\xi + \sum_{k=1}^m c(t_n, x_i, e_k)\right)\right] \\
&= \sum_{m=0}^{\infty} e^{-\lambda\Delta t_n} \frac{(\lambda\Delta t_n)^m}{m!} \mathbb{E}\left[Y^{n+1}\left(x_i + b(t_n, x_i)\Delta t_n + \sigma(t_n, x_i)\sqrt{\Delta t_n}\xi + \sum_{k=1}^m c(t_n, x_i, e_k)\right)\right] \\
&= e^{-\lambda\Delta t_n} \int_{\mathbb{R}^d} Y^{n+1}\left(x_i + b(t_n, x_i)\Delta t_n + \sigma(t_n, x_i)\sqrt{\Delta t_n}\xi\right) \varrho^d(\xi) d\xi \\
&\quad + \sum_{m=1}^{\infty} e^{-\lambda\Delta t_n} \frac{(\lambda\Delta t_n)^m}{m!} \left\{ \int_{\mathbb{R}^d} \int_E \dots \int_E Y^{n+1}\left(x_i + b(t_n, x_i)\Delta t_n \right. \right. \\
&\quad \left. \left. + \sigma(t_n, x_i)\sqrt{\Delta t_n}\xi + \sum_{k=1}^m c(t_n, x_i, e_k)\right) \varrho^d(\xi) \prod_{k=1}^m \rho(e_k) d\xi de_1 \dots de_m \right\}, \tag{5.4}
\end{aligned}$$

where  $c(t_n, x_i, e_k) = c(t_n, x_i, e_k^1, \dots, e_k^q)$  for  $k = 1, \dots, m$  is the size of the  $k$ -th jump and  $\{c(t_n, x_i, e_k)\}_{k=1}^m$  follows the joint distribution  $\prod_{k=1}^m \rho(e_k)$ .

Now we study how to approximate  $\mathbb{E}_{t_n}^{x_i}[Y^{n+1}]$  in (5.4). First, we observe that the probability of having  $m$  jumps within  $(t_n, t_{n+1}]$  is of order  $\mathcal{O}((\lambda\Delta t_n)^m)$ , thus the sum of the infinite sequence in (5.4) can be approximated by the sum of a finite sequence by retaining finite number of jumps. We denote by  $\mathbb{E}_{t_n, M_y}^{x_i}[Y^{n+1}]$  the approximation of  $\mathbb{E}_{t_n}^{x_i}[Y^{n+1}]$  by retaining the first  $M_y$  jumps within  $(t_n, t_{n+1}]$ . Then, it is easy to see that the error introduced by the truncation is of order  $\mathcal{O}((\lambda\Delta t_n)^{M_y+1})$ , so that  $M_y = 2$  is necessary to match the local truncation error introduced by the semi-discrete scheme in (3.25). An analogous notation  $\mathbb{E}_{t_n, M_f}^{x_i}[f^{n+1}]$  is used to represent the approximation of  $\mathbb{E}_{t_n}^{x_i}[f^{n+1}]$  by retaining the first  $M_f$  jumps, where  $M_f = 1$  is sufficient to match the local truncation error in (3.25).

Next, we also need to approximate a  $d$ -dimensional integral with respect to  $\xi$  for  $m = 0$ , and an  $m \times q + d$  dimensional integral with respect to  $(\xi, e_1, \dots, e_m)$  for  $m = 1, \dots, M_y$ . This can be accomplished by selecting an appropriate quadrature rule based on the properties of  $\varrho(\xi)$ ,  $\rho(e)$  and the smoothness of  $Y^{n+1}(x)$  with respect to  $x$ . A straightforward choice is to use Monte Carlo methods by drawing samples from  $\varrho^d(\xi)$  and  $\prod_{k=1}^m \rho(e_k)$ , but they are overall inefficient because of the slow convergence. When  $Y^{n+1}(x)$  is sufficiently smooth with respect to  $x$ , an alternative way is to use the tensor product of high-order one-dimensional quadrature rules, e.g., Newton-Cotes rules and Gaussian rules, etc. For example, the integrals with respect to  $\xi$  in (5.4) can be approximated using the tensor product of the Gauss-Hermite rule [35]. For the integrals with respect to  $(e_1, \dots, e_m)$ , the Gauss-Legendre rule is a good choice when  $\rho(e)$  is compactly supported, e.g.,  $e$  follows a uniform distribution; the Gauss-Laguerre rule is appropriate when  $\rho(e)$  is the density of an exponential distribution. Without loss of generality, for  $m = 0, \dots, M_y$ , we denote by  $\{w_i^m, s_i^m\}_{i=1}^{S_m}$  and  $\{v_j^m, q_j^m\}_{j=1}^{Q_m}$  to represent the chosen quadrature rule for estimating the integrals in (5.4) with respect to  $\xi$  and  $(e_1, \dots, e_m)$ , respectively, where  $\{w_i^m\}_{i=1}^{S_m}$ ,  $\{v_j^m\}_{j=1}^{Q_m}$  are quadrature weights and  $\{s_i^m\}_{i=1}^{S_m}$ ,  $\{q_j^m\}_{j=1}^{Q_m}$  are quadrature points. Note that  $q_j^m$  for  $j = 1, \dots, Q_m$  has  $m$  components, denoted by  $\{q_{j,1}^m, \dots, q_{j,m}^m\}$ , which correspond to the quadrature abscissa for  $(e_1, \dots, e_m)$ . Then, the approximation of  $\mathbb{E}_{t_n}^{x_i}[Y^{n+1}]$ , denoted by  $\hat{\mathbb{E}}_{t_n, M_y}^{x_i}[Y^{n+1}]$ , is represented by

$$\begin{aligned}
\widehat{\mathbb{E}}_{t_n, M_y}^{x_i} [Y^{n+1}] &:= e^{-\lambda \Delta t_n} \sum_{i=1}^{S_0} w_i^0 Y^{n+1} \left( x_i + b(t_n, x_i) \Delta t_n + \sigma(t_n, x_i) \sqrt{\Delta t_n} s_i^0 \right) \\
&+ \sum_{m=1}^{M_y} e^{-\lambda \Delta t_n} \frac{(\lambda \Delta t_n)^m}{m!} \left[ \sum_{i=1}^{S_m} \sum_{j=1}^{Q_m} w_i^m v_j^m Y^{n+1} \left( x_i + b(t_n, x_i) \Delta t_n \right. \right. \\
&\quad \left. \left. + \sigma(t_n, x_i) \sqrt{\Delta t_n} s_i^m + \sum_{k=1}^m c(t_n, x_i, q_{j,k}^m) \right) \right]. \tag{5.5}
\end{aligned}$$

Analogously, the approximation of  $\mathbb{E}_{t_n}^{x_i} [f^{n+1}]$ , denoted by  $\widehat{\mathbb{E}}_{t_n, M_f}^{x_i} [f^{n+1}]$ , can be obtained by replacing  $Y^{n+1}$  with  $f^{n+1}$  in (5.5). For the conditional expectations  $\mathbb{E}_{t_n}^{x_i} [Y^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top]$  and  $\mathbb{E}_{t_n}^{x_i} [f^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top]$ , we observe that each component of  $\Delta \tilde{W}_{t_{n+1}} = (\Delta \tilde{W}_{t_{n+1}}^1, \dots, \Delta \tilde{W}_{t_{n+1}}^d)$  defined in (3.12) can be represented by

$$\Delta \tilde{W}_{t_{n+1}}^i = \frac{\sqrt{\Delta t_n}}{2} (\xi^i + \sqrt{3} \tilde{\xi}^i) \quad \text{for } i = 1, \dots, d,$$

where  $\xi^i, \tilde{\xi}^i$  are independent random variables following standard normal distribution, and  $\xi^i$  is the same random variable as in (5.3). As such, another  $d$ -dimensional integral with respect to  $\tilde{\xi} = (\tilde{\xi}^1, \dots, \tilde{\xi}^d)$  is needed in (5.4) and (5.5) to define  $\widehat{\mathbb{E}}_{t_n, M_y}^{x_i} [Y^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top]$  and  $\widehat{\mathbb{E}}_{t_n, M_f}^{x_i} [f^{n+1} \Delta \tilde{W}_{t_{n+1}}^\top]$ . For the conditional expectations  $\mathbb{E}_{t_n}^{x_i} [Y^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*]$  and  $\mathbb{E}_{t_n}^{x_i} [f^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*]$ , we can see that  $\Delta \tilde{\mu}_{t_{n+1}}^*$  defined in (3.17) can be represented by

$$\begin{aligned}
\Delta \tilde{\mu}_{t_{n+1}}^* &= \int_{t_n}^{t_{n+1}} \int_E \left( 2 - \frac{3(t - t_n)}{\Delta t_n} \right) \eta(e) [\mu(de, dt) - \lambda(de)dt] \\
&= \sum_{k=N_{t_n}+1}^{N_{t_{n+1}}} \left( 2 - \frac{3(\tau_k - t_n)}{\Delta t_n} \right) \eta(e_k) - \frac{\lambda \Delta t_n}{2} \int_E \eta(e) \rho(e) de, \tag{5.6}
\end{aligned}$$

where  $N_t$  for  $t \in [t_n, t_{n+1}]$  is the standard Poisson process and  $\tau_k$  for  $k = N_{t_n}+1, \dots, N_{t_{n+1}}$  is the jump time instant of the  $k$ -th jump within  $(t_n, t_{n+1}]$ . Hence,  $\mathbb{E}_{t_n}^{x_i} [Y^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*]$  involves another integral with respect to  $\tau_k$  compared to  $\mathbb{E}_{t_n}^{x_i} [Y^{n+1}]$ , which requires an additional quadrature rule in (5.5) to construct  $\widehat{\mathbb{E}}_{t_n, M_y}^{x_i} [Y^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*]$ .

Based on the quadrature rules used in (5.5), we observe that it is highly possible the quadrature points do not belong to the spatial grid  $\mathcal{S}$ . In this case, we follow the same strategy as in [33, 35] to resolve this issue, i.e., constructing piecewise Lagrange interpolating polynomials based on  $\mathcal{S}$  to interpolate the integrands at non-grid quadrature points. Again, taking  $Y^{n+1}(x)$  as an example, it can be approximated by

$$Y^{n+1}(x) \approx \widehat{Y}^{n+1}(x) := \sum_{j_1=1}^{p+1} \dots \sum_{j_q=1}^{p+1} \left[ Y_{(i_{j_1}, \dots, i_{j_q})}^{n+1} \prod_{k=1}^q \prod_{\substack{1 \leq j \leq p+1 \\ j \neq j_k}} \frac{x^k - x_{i_j}^k}{x_{i_{j_k}}^k - x_{i_j}^k} \right],$$

where  $\widehat{Y}^{n+1}(x)$  is a  $p$ -th order tensor-product Lagrange interpolating polynomial and  $Y_{(i_{j_1}, \dots, i_{j_q})}^{n+1}$  is the approximate solution of  $Y^{n+1}(x)$  at the spatial point  $(x_{i_{j_1}}^1, \dots, x_{i_{j_q}}^q)$ . For  $k = 1, \dots, q$ , the interpolation points  $\{x_{i_j}^k\}_{j=1}^{p+1} \subset \mathcal{S}^k$  are the *closest*  $p+1$  neighboring points of  $x^k$ , such that  $(x_{i_{j_1}}^1, \dots, x_{i_{j_q}}^q)$  for  $j_k = 1, \dots, p+1$  and  $k = 1, \dots, q$  constitute a local tensor-product sub-grid

around  $x$ . In summary, the fully discrete scheme of the FBSDEs in (2.1) is given as follows:

**Scheme 2.** Given initial condition  $X_0$  for the forward SDE in (2.1) and the terminal condition  $\varphi(X_T)$  for the backward SDE in (2.1), solve the approximate solution  $(Y_{\mathbf{i}}^n, Z_{\mathbf{i}}^n, \Gamma_{\mathbf{i}}^n)$ , for  $n = N-1, \dots, 0$  and  $\mathbf{i} \in \mathbb{Z}^q$ , by

$$X^{n+1} = x_{\mathbf{i}} + \Phi(t_n, t_{n+1}, x_{\mathbf{i}}, I_{\mathcal{J} \in \mathcal{A}_{\beta}}), \quad (5.7)$$

$$Y_{\mathbf{i}}^n = \widehat{\mathbb{E}}_{t_n, M_y}^{x_{\mathbf{i}}} [\widehat{Y}^{n+1}] + \frac{1}{2} \Delta t_n f_{\mathbf{i}}^n + \frac{1}{2} \Delta t_n \widehat{\mathbb{E}}_{t_n, M_f}^{x_{\mathbf{i}}} [\widehat{f}^{n+1}], \quad (5.8)$$

$$\frac{1}{2} \Delta t_n Z_{\mathbf{i}}^n = \widehat{\mathbb{E}}_{t_n, M_y}^{x_{\mathbf{i}}} [\widehat{Y}^{n+1} \Delta \tilde{W}_{t_{n+1}}^{\top}] + \Delta t_n \widehat{\mathbb{E}}_{t_n, M_f}^{x_{\mathbf{i}}} [\widehat{f}^{n+1} \Delta \tilde{W}_{t_{n+1}}^{\top}], \quad (5.9)$$

$$\frac{1}{2} \Delta t_n \Gamma_{\mathbf{i}}^n = \widehat{\mathbb{E}}_{t_n, M_y}^{x_{\mathbf{i}}} [\widehat{Y}^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*] + \Delta t_n \widehat{\mathbb{E}}_{t_n, M_f}^{x_{\mathbf{i}}} [\widehat{f}^{n+1} \Delta \tilde{\mu}_{t_{n+1}}^*], \quad (5.10)$$

where  $\widehat{f}^{n+1} = f(t_{n+1}, X^{n+1}, \widehat{Y}^{n+1}, \widehat{Z}^{n+1}, \widehat{\Gamma}^{n+1})$ ,  $f_{\mathbf{i}}^n = (t_n, x_{\mathbf{i}}, Y_{\mathbf{i}}^n, Z_{\mathbf{i}}^n, \Gamma_{\mathbf{i}}^n)$ ,  $\Delta \tilde{W}_{t_{n+1}}^{\top}$  and  $\Delta \tilde{\mu}_{t_{n+1}}^*$  are defined in (3.12) and (3.17) with  $s = t_{n+1}$ , respectively.

Similar to the semi-discrete scheme, Scheme 2 can be directly used as a fully discrete scheme for the PIDE in (2.2). The solution  $u(t_n, x_{\mathbf{i}})$  of the PIDE is approximated by  $Y_{\mathbf{i}}^n$  for  $n = 0, \dots, N-1$  and  $\mathbf{i} \in \mathbb{Z}^q$ . We observe that at each grid point  $(t_n, x_{\mathbf{i}})$ , the computation of  $Y_{\mathbf{i}}^n$  only depends on  $(X^{n+1}, Y^{n+1}, Z^{n+1}, \Gamma^{n+1})$  even though an implicit time-stepping scheme is used. This means  $\{Y_{\mathbf{i}}^n\}_{\mathbf{i} \in \mathbb{Z}^q}$  at each time step can be computed *independently*, so that the difficulty of solving linear systems with possibly *dense* matrices, due to the nonlocality of the integral operator, is completely avoided. This feature makes it straightforward to develop massively parallel algorithms and incorporate adaptive spatial interpolation methods.

**Remark 2.** It is noted that the total computational cost of the Scheme 2 is dominated by the cost of approximating  $\widehat{\mathbb{E}}_{t_n}^{x_{\mathbf{i}}}[\cdot]$  at each grid point  $(t_n, x_{\mathbf{i}}) \in \mathcal{T} \times \mathcal{S}$  using the formula in (5.5). For example, when solving a three-dimensional problem  $q = d = 3$  and retaining two Lévy jumps  $M_y = M_f = 2$ , we are facing a large amount of six-dimensional integration problems. In this case, sparse-grid quadrature rules [6, 16, 30] can be used to alleviate the explosion of computational cost due to curse of dimensionality.

## 6. Numerical examples

In this section, we report on the results of two one-dimensional numerical examples that illustrate the accuracy and the effectiveness of Schemes 1 and 2. We take uniform partitions in both temporal and spatial domains with the time and space step sizes denoted by  $\Delta t$  and  $\Delta x$ , respectively. The time step number  $N$  is then given by  $N = T/\Delta t$  where  $T$  is the terminal time. For the sake of illustration, we only solve FBSDEs on bounded spatial domains. The goal is to test the convergence rates of time discretization and spatial interpolation with respect to  $\Delta t$  and  $\Delta x$ , respectively. To this end, we always set the number of quadrature points to be sufficiently large, so that the error contributed by the use of quadrature rules is too small to affect the convergence rates of interest.



### 6.1. Example 1

We consider the following *nonlinear* FBSDEs:

$$\begin{cases} dX_t = dW_t + \int_E e \tilde{\mu}(de, dt), \\ -dY_t = \left\{ \frac{(Y_t - 2) \cdot \exp(Y_t)}{2 \exp[\sin(X_t + t) + 2]} - \frac{Z_t Y_t}{\sin(X_t + t) + 2} - \Gamma_t \right\} dt - Z_t dW_t - \int_E U(e) \tilde{\mu}(de, dt), \end{cases} \quad (6.1)$$

where the terminal condition is  $\varphi(X_T) = \sin(X_T + T) + 2$ . The Lèvy measure is defined by

$$\lambda(de) = \lambda \rho(e) de := \mathcal{X}_{[-\delta, \delta]}(e) de \quad \text{with } \delta > 0, \quad (6.2)$$

where  $\mathcal{X}_{[-\delta, \delta]}(e)$  is the characteristic function of the interval  $[-\delta, \delta]$ , so that  $\lambda = 2\delta$  is the jump intensity and  $\rho(e) = \frac{1}{2\delta} \mathcal{X}_{[-\delta, \delta]}(e)$  is the density function of a uniform distribution on  $[-\delta, \delta]$ . The exact solution of the FBSDEs is

$$\begin{cases} Y_t = \sin(X_t + t) + 2, \\ Z_t = \cos(X_t + t), \\ \Gamma_t = \cos(X_t + t - \delta) - \cos(X_t + t + \delta) - 2\delta \sin(X_t + t). \end{cases} \quad (6.3)$$

Accordingly, the PIDE corresponding to (6.1) is

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{1}{2} \frac{\partial^2 u}{\partial x^2} + \int_E (u(t, x + e) - u(t, x)) \lambda(de) \\ \quad + \frac{(u - 2) \cdot \exp(u)}{2 \exp[\sin(x + t) + 2]} - \frac{u}{\sin(x + t) + 2} \frac{\partial u}{\partial x} - \mathcal{B}[u] = 0, \\ u(T, x) = \sin(x + T) + 2, \end{cases}$$

where  $\mathcal{B}[u] = \cos(x + t - \delta) - \cos(x + t + \delta) - 2\delta \sin(x + t)$ .

Since the density function  $\rho(e)$  is uniform with the support  $[-\delta, \delta]$ , we use the tensor product of the 8-point Gauss-Legendre rule and the 8-point Gauss-Hermite rule to approximate the integrals involved in  $\mathbb{E}_{t_n}^{x_i}[\cdot]$ .

First, we test the convergence rate with respect to  $\Delta t$  where the terminal time is  $T = 1$ . To this end, we set  $\Delta x = 0.01$  and use piecewise cubic Lagrange interpolation to construct  $\hat{Y}^{n+1}(x)$  for  $n = 0, \dots, N - 1$ , such that the time discretization error dominates the total error. Setting  $\delta = 1$  and  $\Delta t = 2^{-4}, 2^{-5}, 2^{-6}, 2^{-7}, 2^{-8}$ , the numerical results are shown in Table 6.1. As expected, the convergence rate with respect to  $\Delta t$  depends on the number of jumps retained in  $\hat{\mathbb{E}}_{t_n, M_y}^{x_i}[\cdot]$  and  $\hat{\mathbb{E}}_{t_n, M_f}^{x_i}[\cdot]$ . For example, when  $M_y = M_f = 0$ , i.e., no jump is included, our scheme fails to converge. In order to achieve second-order convergence, we must set  $M_y \geq 2$  and  $M_f \geq 1$ .

Next, we test the convergence rate with respect to  $\Delta x$  by setting  $\delta = 1$ ,  $T = 1$ ,  $N = 1024$ ,  $M_y = 3$ ,  $M_f = 2$ , and  $\Delta x = 2^{-2}, 2^{-3}, 2^{-4}, 2^{-5}, 2^{-6}$ . The error is measured in  $L^\infty$  norm. In Table 6.2, we can see that the spatial discretization error decays as expected, i.e., second-order and third-order convergence rates for piecewise linear and piecewise quadratic interpolations, respectively.

Table 6.1: Errors and convergence rates with respect to  $\Delta t$  in Example 1, where  $T = 1$ ,  $\delta = 1$ ,  $\Delta x = 0.01$ , and piecewise cubic Lagrange interpolation is used.

$\ Y_0^{0,x} - \tilde{Y}^0(x)\ _{L^\infty([0,1])}$						
	$\Delta t = 2^{-4}$	$\Delta t = 2^{-5}$	$\Delta t = 2^{-6}$	$\Delta t = 2^{-7}$	$\Delta t = 2^{-8}$	CR
$M_y = 0, M_f = 0$	5.178E-1	4.786E-1	4.498E-1	4.290E-1	4.090E-1	0.084
$M_y = 1, M_f = 0$	4.155E-2	1.778E-2	7.855E-3	3.631E-3	1.691E-3	1.153
$M_y = 2, M_f = 1$	4.539E-3	9.878E-4	2.211E-4	5.065E-5	1.144E-5	2.155
$M_y = 3, M_f = 2$	3.414E-3	7.305E-4	1.609E-4	3.646E-5	8.087E-6	2.177
$\ Z_0^{0,x} - \tilde{Z}^0(x)\ _{L^\infty([0,1])}$						
	$\Delta t = 2^{-4}$	$\Delta t = 2^{-5}$	$\Delta t = 2^{-6}$	$\Delta t = 2^{-7}$	$\Delta t = 2^{-8}$	CR
$M_y = 0, M_f = 0$	1.591E-0	2.155E-0	2.505E-0	2.237E-0	2.534E-0	-0.140
$M_y = 1, M_f = 0$	1.629E-1	9.244E-2	5.028E-2	2.156E-2	1.146E-2	0.976
$M_y = 2, M_f = 1$	1.846E-2	5.459E-3	1.536E-3	3.293E-4	8.706E-5	1.951
$M_y = 3, M_f = 2$	1.475E-2	4.230E-3	1.161E-3	2.450E-4	6.376E-5	1.982
$\ \Gamma_0^{0,x} - \tilde{\Gamma}^0(x)\ _{L^\infty([0,1])}$						
	$\Delta t = 2^{-4}$	$\Delta t = 2^{-5}$	$\Delta t = 2^{-6}$	$\Delta t = 2^{-7}$	$\Delta t = 2^{-8}$	CR
$M_y = 0, M_f = 0$	5.165E-1	7.206E-1	6.360E-1	5.444E-1	5.270E-1	0.035
$M_y = 1, M_f = 0$	3.519E-1	1.805E-1	9.636E-2	4.416E-2	2.240E-2	0.998
$M_y = 2, M_f = 1$	2.151E-2	5.364E-3	1.453E-3	3.365E-4	8.458E-5	1.998
$M_y = 3, M_f = 2$	1.549E-2	3.856E-3	1.057E-3	2.467E-4	6.212E-5	1.989

Table 6.2: Errors and convergence rates with respect to  $\Delta x$  in Example 1, where  $T = 1$ ,  $\delta = 1$ ,  $x \in [0, 1]$ ,  $N = 1024$ ,  $M_y = 3$  and  $M_f = 2$ .

Linear interpolation						
	$\Delta x = 2^{-2}$	$\Delta x = 2^{-3}$	$\Delta x = 2^{-4}$	$\Delta x = 2^{-5}$	$\Delta x = 2^{-6}$	CR
$\ Y_0^{0,x} - \tilde{Y}^0(x)\ _\infty$	1.756E-2	6.094E-3	1.968E-3	3.194E-4	1.012E-4	2.036
$\ Z_0^{0,x} - \tilde{Z}^0(x)\ _\infty$	9.273E-2	3.163E-2	8.457E-3	2.278E-3	5.203E-4	1.885
$\ \Gamma_0^{0,x} - \tilde{\Gamma}^0(x)\ _\infty$	2.000E-2	6.812E-3	2.173E-3	3.360E-4	1.081E-4	2.063
Quadratic interpolation						
	$\Delta x = 2^{-2}$	$\Delta x = 2^{-3}$	$\Delta x = 2^{-4}$	$\Delta x = 2^{-5}$	$\Delta x = 2^{-6}$	CR
$\ Y_0^{0,x} - \tilde{Y}^0(x)\ _\infty$	6.183E-2	8.004E-3	9.941E-4	1.266E-4	1.438E-5	3.012
$\ Z_0^{0,x} - \tilde{Z}^0(x)\ _\infty$	5.537E-2	7.016E-3	1.583E-3	1.266E-4	1.096E-5	3.039
$\ \Gamma_0^{0,x} - \tilde{\Gamma}^0(x)\ _\infty$	1.927E-2	3.489E-3	3.527E-4	5.531E-5	5.587E-6	2.948

## 6.2. Example 2

We consider the following nonlinear FBSDE:

$$\left\{ \begin{array}{l} dX_t = \sin(2X_t + t)dt + [\cos(X_t) + t + 2]dW_t + \int_E e \tilde{\mu}(de, dt), \\ -dY_t = \left\{ -\frac{\sin(2X_t + t)Y_t Z_t}{[\cos(X_t) + t + 2](\sin(t) + 2)\exp(-X_t)} \right. \\ \quad \left. - 0.5[\cos(X_t) + t + 2]^2 Y_t - \Gamma_t \right\} dt - Z_t dW_t - \int_E U(e) \tilde{\mu}(de, dt), \end{array} \right. \quad (6.4)$$

where the terminal condition is  $\varphi(X_T) = [\sin(T) + 2]\exp(-X_T)$ . The Lèvy measure  $\tilde{\mu}(de, dt)$  is defined as in (6.2). The exact solution of the FBSDEs is

$$\left\{ \begin{array}{l} Y_t = (\sin(t) + 2)\exp(-X_t), \\ Z_t = -(\cos(X_t) + t + 2)(\sin(t) + 2)\exp(-X_t), \\ \Gamma_t = (\sin(t) + 2)[\exp(-X_t + \delta) - \exp(-X_t - \delta) - 2\delta\exp(-X_t)]. \end{array} \right. \quad (6.5)$$

Accordingly, the PIDE corresponding to (6.4) is

$$\begin{cases} \frac{\partial u}{\partial t} + \sin(2x+t) \frac{\partial u}{\partial x} + \frac{1}{2} [\cos(x) + t + 2]^2 \frac{\partial^2 u}{\partial x^2} \\ + \int_E (u(t, x+e) - u(t, x)) \lambda(de) - \frac{\sin(2x+t)u(t, x)}{(\sin(t) + 2) \exp(-x)} \frac{\partial u}{\partial x} \\ - \frac{1}{2} [\cos(x) + t + 2]^2 u(t, x) - \mathcal{B}[u] = 0, \\ u(T, x) = \sin(T+2) \exp(-x), \end{cases}$$

where  $\mathcal{B}[u] = (\sin(t) + 2)[\exp(-x + \delta) - \exp(-x - \delta) - 2\delta \exp(-x)]$ . Similar to Example 1, we use the tensor product of the 8-point Gauss-Legendre rule and the 8-point Gauss-Hermite rule to approximate the integrals involved in  $\mathbb{E}_{t_n}^{x_i}[\cdot]$ . In this example, we use forward Euler scheme in (5.2), such that the overall convergence rate will be expected to be first order.

First, we test the convergence rate with respect to  $\Delta t$  where the terminal time is  $T = 1$ . To this end, we set  $\Delta x = 0.01$  and use piecewise cubic Lagrange interpolation to construct  $\hat{Y}^{n+1}(x)$  for  $n = 0, \dots, N-1$ , so that the time discretization error dominates the total error. Setting  $\delta = 1$  and  $\Delta t = 2^{-5}, 2^{-6}, 2^{-7}, 2^{-8}, 2^{-9}$ , the numerical results are shown in Table 6.3. As expected, the convergence rate with respect to  $\Delta t$  depends on the number of jumps retained in constructing  $\hat{\mathbb{E}}_{t_n, M_y}^{x_i}[\cdot]$  and  $\hat{\mathbb{E}}_{t_n, M_f}^{x_i}[\cdot]$ . In this case, we can only achieve, at most, first-order convergence with respect to  $\Delta t$  due to the use of the forward Euler scheme.

Next, we test the convergence rate with respect to  $\Delta x$  by setting  $\delta = 1$ ,  $T = 1$ ,  $N = 1024$ ,  $M_y = 2$  and  $M_f = 1$ . The spatial mesh size is set to  $\Delta x = 2^{-5}, 2^{-6}, 2^{-7}, 2^{-8}, 2^{-9}$  for linear interpolation and  $\Delta x = 2^{-2}, 2^{-3}, 2^{-4}, 2^{-5}, 2^{-6}$  for quadratic interpolation. The error is measured in  $L^\infty$  norm. In Table 6.4, we can see that the spatial discretization error decays as expected, i.e., second-order and third-order convergence rates for piecewise linear and piecewise quadratic interpolations, respectively.

Table 6.3: Errors and convergence rates with respect to  $\Delta t$  in Example 2, where  $T = 1$ ,  $\delta = 1$ ,  $\Delta x = 0.01$ , and piecewise cubic Lagrange interpolation are used.

$\ Y_0^{0,x} - \bar{Y}^0(x)\ _{L^\infty([0,1])}$						
	$\Delta t = 2^{-5}$	$\Delta t = 2^{-6}$	$\Delta t = 2^{-7}$	$\Delta t = 2^{-8}$	$\Delta t = 2^{-9}$	CR
$M_y = 0, M_f = 0$	1.331E-1	9.191E-2	7.689E-2	6.763E-2	6.206E-2	0.264
$M_y = 1, M_f = 0$	2.996E-2	1.362E-2	5.339E-3	2.226E-3	1.998E-3	1.043
$M_y = 2, M_f = 1$	3.071E-2	1.047E-2	3.835E-3	1.613E-3	7.155E-4	1.355
$\ Z_0^{0,x} - \bar{Z}^0(x)\ _{L^\infty([0,1])}$						
	$\Delta t = 2^{-5}$	$\Delta t = 2^{-6}$	$\Delta t = 2^{-7}$	$\Delta t = 2^{-8}$	$\Delta t = 2^{-9}$	CR
$M_y = 0, M_f = 0$	4.757E-1	4.522E-1	5.558E-1	6.583E-1	6.881E-1	-0.161
$M_y = 1, M_f = 0$	3.779E-1	1.733E-1	7.756E-2	3.598E-2	1.661E-2	1.128
$M_y = 2, M_f = 1$	1.129E-1	5.351E-2	2.714E-2	1.201E-2	5.771E-3	1.074
$\ \Gamma_0^{0,x} - \bar{\Gamma}^0(x)\ _{L^\infty([0,1])}$						
	$\Delta t = 2^{-5}$	$\Delta t = 2^{-6}$	$\Delta t = 2^{-7}$	$\Delta t = 2^{-8}$	$\Delta t = 2^{-9}$	CR
$M_y = 0, M_f = 0$	1.379E-1	1.194E-1	8.948E-2	8.391E-2	7.765E-2	0.217
$M_y = 1, M_f = 0$	5.125E-2	2.528E-2	1.226E-2	7.598E-3	3.996E-3	0.909
$M_y = 2, M_f = 1$	4.789E-2	2.116E-2	1.057E-2	5.373E-3	2.509E-3	1.049

## 7. Concluding remarks

In this work, we propose new numerical schemes for decoupled forward-backward stochastic differential equations with jumps, which feature high-order temporal and spatial convergence

Table 6.4: Errors and convergence rates with respect to  $\Delta x$  in Example 2, where  $T = 1$ ,  $\delta = 1$ ,  $x \in [0, 1]$ ,  $N = 1024$ ,  $M_y = 2$  and  $M_f = 1$ .

Linear interpolation						
	$\Delta x = 2^{-5}$	$\Delta x = 2^{-6}$	$\Delta x = 2^{-7}$	$\Delta x = 2^{-8}$	$\Delta x = 2^{-9}$	CR
$\ Y_0^{0,x} - \hat{Y}^0(x)\ _\infty$	2.981E-3	8.383E-4	2.227E-4	5.583E-5	1.652E-5	1.899
$\ Z_0^{0,x} - \hat{Z}^0(x)\ _\infty$	8.538E-2	2.774E-2	6.109E-3	1.810E-3	4.206E-4	1.926
$\ \Gamma_0^{0,x} - \hat{\Gamma}^0(x)\ _\infty$	5.859E-3	1.537E-3	2.667E-4	7.729E-5	2.097E-5	2.056
Quadratic interpolation						
	$\Delta x = 2^{-2}$	$\Delta x = 2^{-3}$	$\Delta x = 2^{-4}$	$\Delta x = 2^{-5}$	$\Delta x = 2^{-6}$	CR
$\ Y_0^{0,x} - \hat{Y}^0(x)\ _\infty$	3.127E-2	3.916E-3	6.021E-4	8.410E-5	1.093E-5	2.856
$\ Z_0^{0,x} - \hat{Z}^0(x)\ _\infty$	8.945E-2	1.158E-2	1.376E-3	1.966E-4	2.638E-5	2.935
$\ \Gamma_0^{0,x} - \hat{\Gamma}^0(x)\ _\infty$	7.512E-3	1.349E-3	2.109E-4	3.236E-5	2.807E-6	2.815

rates. This advantage has been verified by both theoretical analysis and numerical experiments. Meanwhile, we also realized that our schemes cannot achieve the desired convergence rates in the sense that the solution of the FBSDEs does not satisfy the necessary regularity conditions. For example, this may happen in real-world financial problems, such as option pricing. However, the regularity conditions do not limit the applicability of the proposed approach, because our method can be directly employed as a probabilistic scheme for related PIDEs which are widely used to describe anomalous transport in subsurface flow and plasma physics. In these settings, there is a variety of problems satisfying the regularity conditions, and high-order schemes are highly desired. Moreover, compared to existing deterministic approaches (e.g., finite elements) for the PIDEs, the ability to completely avoids the solution of *dense* linear systems, as well as to utilize efficient adaptive approximation, and the potential of massively parallel implementation, make our technique highly advantageous. Our future works will focus on extending the proposed numerical schemes to the case of Poisson random measures with *infinite* activities, and integrating sparse grid methods for high-dimensional FBSDEs with jumps.

**Acknowledgments.** This work is partially supported by the National Natural Science Foundations of China under grant numbers 91130003 and 11171189; by Natural Science Foundation of Shandong Province under grant number ZR2011AZ002; by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Applied Mathematics program under contract number ERKJE45; and by the Laboratory Directed Research and Development program at the Oak Ridge National Laboratory, which is operated by UT-Battelle, LLC, for the U.S. Department of Energy under Contract DE-AC05-00OR22725.

## References

- [1] Guy Barles, Rainer Buckdahn, and Etienne Pardoux, *Backward stochastic differential equations and integral-partial differential equations*, Stochastics and Stochastics Reports **60** (1997), no. 1-2, 57–83.
- [2] Dirk Becherer, *Bounded solutions to backward SDEs with jumps for utility optimization and indifference hedging*, The Annals of Applied Probability **16** (2006), no. 4, 2027–2054.
- [3] Christian Bender and Robert Denk, *A forward scheme for backward SDEs*, Stochastic Processes and their Applications **117** (2007), no. 12, 1793–1812.
- [4] Bruno Bouchard and Romuald Elie, *Discrete-time approximation of decoupled Forward–Backward SDE with jumps*, Stochastic Processes and their Applications **118** (2008), no. 1, 53–75.
- [5] Bruno Bouchard, Romuald Elie, and Nizar Touzi, *Discrete-time approximation of BSDEs and probabilistic schemes for fully nonlinear PDEs*, Advanced financial modelling, Walter de Gruyter,

- Berlin, 2009, pp. 91–124.
- [6] Hans-Joachim Bungartz and Michael Griebel, *Sparse grids*, Acta Numerica **13** (2004), 1–123.
  - [7] Jean-Francois Chassagneux and Dan Crisan, *Runge–Kutta schemes for backward stochastic differential equations*, The Annals of Applied Probability **24** (2014), no. 2, 679–720.
  - [8] Chuchu Chen and Jia Hong, *Mean-square convergence of numerical approximations for a class of backward stochastic differential equations*, Discrete and Continuous Dynamical Systems-Series B, **8** (2013), 2051–2067.
  - [9] S Crépey, *Financial Modeling, A Backward Stochastic Differential Equations Perspective*, Springer Finance, 2013.
  - [10] Dan Crisan and Konstantinos Manolarakis, *Second order discretization of backward SDEs and simulation with the cubature method*, The Annals of Applied Probability **24** (2014), no. 2, 652–678.
  - [11] Lukasz Delong, *Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications*, BSDEs with Jumps, Springer Science & Business, June 2013.
  - [12] Jim Douglas, Jin Ma, and Philip Protter, *Numerical methods for forward-backward stochastic differential equations*, The Annals of Applied Probability **6** (1996), no. 3, 940–968.
  - [13] Nicole EL Karoui, Shige Peng, and M C Quenez, *Backward stochastic differential equations in finance*, Mathematical Finance. An International Journal of Mathematics, Statistics and Financial Economics **7** (1997), no. 1, 1–71.
  - [14] Anne Eyraud-Loisel, *Backward stochastic differential equations with enlarged filtration: Option hedging of an insider trader in a financial market with jumps*, Stochastic Processes and their Applications **115** (2005), no. 11, 1745–1763.
  - [15] Emmanuel Gobet, Jean-Philippe Lemor, and Xavier Warin, *A regression-based Monte Carlo method to solve backward stochastic differential equations*, The Annals of Applied Probability **15** (2005), no. 3, 2172–2202.
  - [16] Max D Gunzburger, Clayton G Webster, and Guannan Zhang, *Stochastic finite element methods for partial differential equations with random input data*, Acta Numerica **23** (2014), 521–650.
  - [17] Desmond J Higham and Peter E Kloeden, *Numerical methods for nonlinear stochastic differential equations with jumps*, Numerische Mathematik **101** (2005), no. 1, 101–119.
  - [18] Antoine Lejay, Ernesto Mordecki, Soledad Torres, *Numerical approximation of Backward Stochastic Differential Equations with Jumps*, (2007).
  - [19] Ralf Metzler and Joseph Klafter, *The random walk’s guide to anomalous diffusion: a fractional dynamics approach*, Physics Reports **339** (2000), no. 1, 1–77.
  - [20] Bernt Øksendal, Agnès Sulem, and Agnès Sulem, *Maximum Principles for Optimal Control of Forward-Backward Stochastic Differential Equations with Jumps*, SIAM Journal on control and optimization **48** (2010), no. 5, 2945–2976.
  - [21] Etienne Pardoux and ShigePeng, *Backward stochastic differential equations and quasilinear parabolic partial differential equations*, Stochastic Partial Differential Equations and Their Applications, Springer Berlin Heidelberg, Berlin/Heidelberg, January 1992, pp. 200–217.
  - [22] Etienne Pardoux and Shige Peng, *Adapted solution of a backward stochastic differential equation*, Systems & Control Letters **14** (1990), no. 1, 55–61.
  - [23] Shige Peng, *Backward stochastic differential equations and applications to optimal control*, Applied Mathematics and Optimization **27** (1993), no. 2, 125–144.
  - [24] Shige Peng,, *Nonlinear Expectations, Nonlinear Evaluations and Risk Measures*, Stochastic Methods in Finance, Springer Berlin Heidelberg, Berlin, Heidelberg, January 2004, pp. 165–253.
  - [25] Eckhard Platen and Nicola Bruti-Liberati, *Numerical Solution of Stochastic Differential Equations with Jumps in Finance*, Stochastic Modelling and Applied Probability, vol. 64, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.
  - [26] Marie-Claire Quenez and Agnès Sulem, *BSDEs with jumps, optimization and applications to dynamic risk measures*, Stochastic Processes and their Applications **123** (2013), no. 8, 3328–3357.

- [27] Situ Rong, *On solutions of backward stochastic differential equations with jumps and applications*, Stochastic Processes and their Applications **66** (1997), no. 2, 209–236.
- [28] Manuela Royer, *Backward stochastic differential equations with jumps and related non-linear expectations*, Stochastic Processes and their Applications **116** (2006), no. 10, 1358–1376.
- [29] Shanjian Tang and Xunjing Li, *Necessary Conditions for Optimal Control of Stochastic Systems with Random Jumps*, SIAM Journal on control and optimization **32** (1994), no. 5, 1447–1475.
- [30] Guannan Zhang, Max Gunzburger, and Weidong Zhao, *A sparse-grid method for multi-dimensional backward stochastic differential equations*, Journal of Computational Mathematics **31** (2013), no. 3, 221–248.
- [31] Guannan Zhang, Weidong Zhao, Clayton Webster, and Max Gunzburger, *A numerical scheme for nonlocal diffusion equations via backward stochastic differential equations with jumps*, ORNL technical report 2014/449 (2014).
- [32] Jianfeng Zhang, *A numerical scheme for BSDEs*, The Annals of Applied Probability **14** (2004), no. 1, 459–488.
- [33] Weidong Zhao, Lifeng Chen, and Shige Peng, *A New Kind of Accurate Numerical Method for Backward Stochastic Differential Equations*, SIAM Journal on Scientific Computing **28** (2006), no. 4, 1563–1581.
- [34] Weidong Zhao, Yu Fu, and Tao Zhou, *New Kinds of High-Order Multistep Schemes for Coupled Forward Backward Stochastic Differential Equations*, SIAM Journal on Scientific Computing **36** (2014), no. 4, A1731–A1751.
- [35] Weidong Zhao, Guannan Zhang, and Lili Ju, *A stable multistep scheme for solving backward stochastic differential equations*, SIAM Journal on Numerical Analysis **48** (2010), no. 4, 1369–1394.
- [36] Weidong Zhao, Wei Zhang, and Lili Ju, *A numerical method and its error estimates for the decoupled forward-backward stochastic differential equations*, Commun Comput Phys **15** (2014), 618–646.